CATALYSIS BY DNA ENZYMES: ROLES OF LANTHANIDE(III) IONS AND DNA APTAMER MODULES

BY

VICTOR DOKUKIN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor Scott K. Silverman, Chair
Professor Gregory S. Girolami
Assistant Professor Douglas A. Mitchell
Professor Wilfred A. van der Donk

Abstract

For many years scientists considered DNA and RNA to be merely the means for the storage and transmission of genetic information. Catalytic roles of some RNA molecules, called ribozymes, were discovered in the early 1980s. Considering structural similarities between RNA and DNA, it was reasonable to evaluate catalytic competence of DNA. Indeed, the first deoxyribozymes were reported in 1994. However, DNA enzymes are not found in nature and can only be identified from a pool of random sequences via an artificial selection process, referred to as the in vitro selection. DNA catalysts capable of supporting various reactivities have been found.

All deoxyribozymes can be classified as metalloenzymes because they require metal ion cofactors, usually at millimolar concentrations. For example, DNA catalysts for single-stranded DNA hydrolysis, previously identified in the Silverman laboratory, are functional only in presence of Mn^{2+} and Zn^{2+} cofactors at millimolar concentrations. Chapter 2 describes the discovery via the selection process of two DNA-hydrolyzing DNA enzymes that require low micromolar concentrations of a lanthanide ion as the sole polyvalent metal ion cofactor. The same selection effort also lead to the identification of a number of DNA-hydrolyzing deoxyribozymes that strictly require low micromolar lanthanide as well as millimolar Zn^{2+} concentrations. These DNA catalysts have a range of lanthanide dependences, including some deoxyribozymes that strongly favor one particular lanthanide ion and others that function well with more than one lanthanide ion.

In addition to its catalytic competency, DNA is capable of forming strong non-Watson-Crick binding interactions with various chemical targets, including small molecules. DNA sequences possessing this ability are referred to as DNA aptamers and

can also be identified via the in vitro selection process. Chapter 3 describes an effort to find DNA aptamers of ATP with a further goal of integrating these sequences into random DNA pools for deoxyribozyme selection. Providing a fixed binding module adjacent to the random region during selection was hypothesized to lead to the emergence of DNA sequences tasked exclusively with catalysis. Previously published and well-studied DNA aptamers of ATP were chosen for this task due to a lack of success in identifying new ATP binders.

Chapter 4 describes selection efforts with an integrated DNA aptamer that lead to the identification of a truly modular deoxyribozyme. This DNA enzyme catalyzes tyrosine phosphorylation via the transfer of a phosphoryl group from ATP, which must be present in solution at concentrations similar to the K_d value of the integrated aptamer. Functional contribution of the binding domain to catalysis was confirmed by mutational analysis of the aptamer sequence as well as by substitution of ATP target molecules with their analogs.

Acknowledgements

I owe my deepest gratitude to those who contributed their time and effort in helping me stay on course during my graduate voyage. First and foremost, I would not be at this stage without the support and guidance from my academic advisor, Dr. Scott K. Silverman. Watching him manage the laboratory has taught me how to lead a scientific team and how to organize my projects. His advice during our regular discussions inspired me to grow both scientifically and personally.

I would like to thank the former Silverman group members who welcomed me on board when I was beginning this adventure. Dr. Amit Sachdeva, Dr. Madhav Chandra, Dr. Adrienne (On Yi) Wong and Dr. Ying Xiao have showed me the ropes of catalytic DNA work. I am also grateful to the current members of the laboratory; friendship, advice and encouragement from Ben, Jagdeesh, Spurti, Shannon, Jimmy, Puzhou, Cong and Peter have put wind in my sails during these years.

I appreciate the time that my thesis committee members, Prof. Wilfred A. van der Donk, Prof. Gregory S. Girolami and Prof. Douglas M. Mitchell have invested in helping me navigate through my graduate career.

I am thankful for the unconditional love and support of my family members. Surrounded by care from my wife, my mother, my grandmother and my brother, I always felt in safe harbor.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction to Nucleic Acids

1.1 Importance of Nucleic Acids

Nucleic acids are essential natural polymeric molecules found in all known forms of life. Biopolymers of DNA (deoxyribonucleic acid) or RNA (ribonucleic acid) are used for genetic storage, encoding, transmission of hereditary information and even as regulatory and catalytic molecules, while monomers are employed in numerous processes as cofactors. In academic laboratories around the world these important molecules continue to be studied, while nucleic acid-based biotechnologies have become a fundamental part of today's innovative biotechnological and biopharmaceutical industries.

1.2 Chemical Structure and Function of DNA and RNA

In general, DNA and RNA, the two types of natural macromolecules, are linear polymers of nucleotides. In turn, each nucleotide consists of three parts (Fig. 1.1): (i) a five-membered pentose sugar ring; (ii) a nucleobase attached to the 1′-carbon of the sugar via a glycosidic bond; (iii) a phosphate group, through which the 3′-carbon of one nucleotide is connected to the 5′-carbon of its neighbor. Only two structural differences exist between DNA and RNA nucleotides: the latter has a hydroxyl group in place of a hydrogen at the 2′-position of the sugar and uses a uracil (U) nucleobase instead of thymine (T). Both polymers have adenine (A), guanine (G) and cytosine (C) as bases. The complementarity of nucleic acid polymers arises from hydrogen-bond interactions between A and T (or U) as well as G and C nucleobases referred to as Watson-Crick base-pairing.

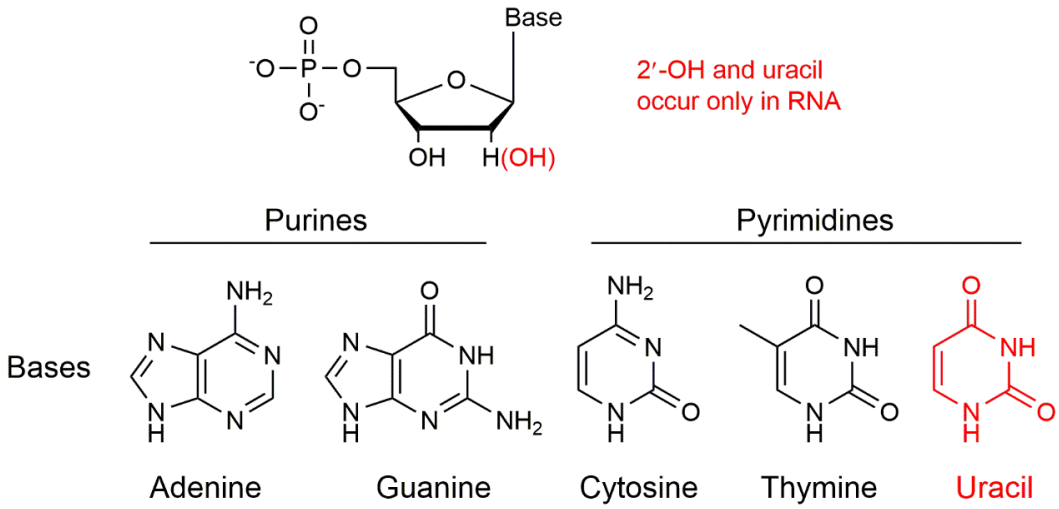

Figure 1.1: A single monomer of DNA or RNA and the five natural nucleobases.

1.2.1 DNA chemical structure and function

DNA in nature is mostly found in a right-handed double-helical form with two DNA strands coiled along the same axis but running in anti-parallel directions; each strand has a pitch of 3.4 nm and a radius of 1.0 nm.[1] There are about 10 nucleotides per full turn of the helix. Two space voids, referred to as grooves, run parallel to the DNA strands. Because the strands are not symmetric with respect to each other, the grooves have different sizes with the major groove 2.2 nm wide and the minor groove 1.2 nm wide.[2] Although hydrogen bonds between complementary nucleobases are important for stability, the major forces defining this structure come from the π–π stacking interactions between highly conjugated purine and pyrimidine rings separated by 0.34 nm. It has been estimated that stacking alone can contribute up to –1.0 kcal/mol per each stacking interaction while base-pairing adds a maximum of –0.8 kcal/mol per base pair.[3]

The Nobel Prize in Physiology or Medicine was awarded to Crick and Watson for their X-ray studies of the DNA structure first published in 1953.[4] The described double-helical form of DNA, also known as the B-form, is its most common composition in living cells. However, other double-helical forms of complementary DNA:DNA, DNA:RNA and RNA:RNA strands are known. One of them, known as the Z form, has a left-handed helical twist.[5] Structures containing three and four DNA strands have also been characterized.[6] Observation of the alternate structures depends on hydration levels, identity and concentration of metal ions, pH and nucleobase make-up, among other factors.

Genes, units of hereditary information of life, are segments of DNA which encode RNA molecules as well as protein sequences, another type of essential macromolecule of life. As such, genes hold the necessary information to build and maintain cells and pass traits of an organism to its offspring. Genes can be as short as 76 nucleotides and as long as 2.2 million nucleotides and are located in chromosomes, structures of highly coiled DNA found inside nuclei of eukaryotic cells and in the cytoplasm of prokaryotes.[7]

1.2.2 RNA chemical structure and function

An important intermediate between DNA in a gene and a protein encoded by that gene is mRNA (messenger RNA). While DNA commonly occurs in nature in a double-helical form, it is typical of RNA to be single-stranded. Although lacking a separate complementary strand, RNA is capable of forming intrastrand hydrogen bonding between stretches of complementary nucleotides, thus folding into complex, three-dimensional structures.

RNA molecules are separated into several categories, based on their function inside a cell. Previously mentioned mRNA serves as an intermediate template for protein synthesis.[8] mRNA is constructed by RNA polymerase via complementary base-pairing from a DNA template located on the antisense strand of a double helix in a process referred to as transcription. mRNA is synthesized in a 5′ to 3′ direction while the DNA strand is read in the opposite, 3′ to 5′ direction. After transcription, mRNA molecules must be processed: noncoding regions are removed during mRNA splicing, the 5′ ends are capped with 7-methylguanosine and a long polyA tails are added to the 3′ end to protect the molecule from degradation.[9] In mRNA, each set of three consecutive nucleotides is called a codon. Four nucleotide identities and three positions in a codon provide for a total of 64 unique codons. Each codon encodes a single natural amino acid with the notable exception of three stop codons each of which terminates protein synthesis. There is a degeneracy of the genetic code as 61 remaining codons specify 20 amino acids; in most cases, codons for the same amino acid differ only in their third position. The information encoded in mRNA is used to build proteins in a process known as translation where the identity of the amino acid to be installed is based on nucleotide base-pairing between sequences of the codon from mRNA and the anticodon in transfer RNA, or tRNA, another type of natural RNA molecule. These highly structured adaptor molecules are encoded by DNA in what is considered to be the set of shortest genes. Single amino acids are attached to the 3′ end of a tRNA molecule. Nucleobase modifications are common throughout a tRNA sequence and most often include methylation. Sometimes the first anticodon nucleotide, also known as a wobble base that is able to base-pair to more than one nucleobase of the codon, is modified to pseudouridine (from U), inosine (from A) or lysidine (from C).[10] The third

most common functional RNA is the ribosomal RNA, or rRNA. Together with structurally required proteins, rRNA makes up the ribosome, the catalyst of the translation process. Ribosome structures vary for prokaryotes and eukaryotes but all can be generalized to having two domains: one large and one small subunit. One eukryotic ribosome is made up of four RNA sequences containing a total of 7216 nucleotides, while that of *E. coli* has three strands and is 4568 bases long.[11] The differences between human and bacterial ribosomes are sufficient enough for specific targeting of the latter by some antibiotics without significant side effects.[12]

1.3 Other Functional RNA Molecules

In addition to RNA molecules that participate in transcriptional and translational processes of gene expression, other functional molecules of this biopolymer are present in organisms. Below are brief descriptions of several such examples: riboswitches, small interfering RNAs and ribozymes.

1.3.1 Riboswitches

In addition to more complicated mechanisms involving macromolecules, small molecules (especially metabolites) can regulate expression of genes which typically code for proteins involved in transport or biogenesis of these metabolites.[13,14] The small molecule targets bind directly to a non-coding regulatory region at the 5′ end of mRNA known as the riboswitch segment. Although they are parts of mRNA molecules,

riboswitches can be discussed separately from simple mRNA due to their unique properties. Even very similar molecular analogs of the small molecule target are discriminated against by riboswitches. This very specific interaction results in a structural change of the sequence due to the presence of regions capable of alternative base-paring near the small molecule binding sites (Fig. 1.2) and regulates expression of a protein it encodes. Most riboswitches downregulate gene expression upon binding of the small-molecule metabolite while a fewer number of these RNAs can turn it on. The evolutionarily conserved segment with such great specificity of binding is called the aptamer region of the riboswitch. The existence of this regulatory mechanism had been postulated in the 1960's but the fact that mRNA can bind metabolites was first shown in 2002 in Ronald Breaker's laboratory at Yale.[15] Prior to that, in 1998, Tina Henkin and coworkers from the Ohio State University first observed this gene regulatory pathway.[16] Interestingly, these naturally occurring polynucleotide binders were found more than a decade after artificial DNA and RNA aptamers were identified.[17,18] Currently known riboswitches can sense various molecules including purines, protein coenzymes, amino acids, metal ions and even negatively charged fluoride ions.[19] Most notable examples of riboswitch regulatory mechanisms involve formation of mutually exclusive conformations of RNA that affect either transcription or translation processes in bacteria. Intrinsic terminator or antiterminator stem-loops are created to inhibit or to activate transcription, respectively. In a similar manner, translation is regulated by the sequestration or release of ribosome-binding sites on mRNA. In eukaryotes, riboswitches regulate gene expression via alternative splicing, which is a mechanism in which exon regions are rejoined in various combinations after removal of introns to produce different mRNA sequences.[13] Binding of

a target molecule by the riboswitch exposes previously hidden splice site. This binding-induced splicing event then leads to either premature termination of the translation process or production of a non-functional polypeptide. The topic of riboswitches is currently very prominent in the field of nucleic acids research as scientists explore the potential for these RNA sequences in applications for medicine and biotechnology.[13]

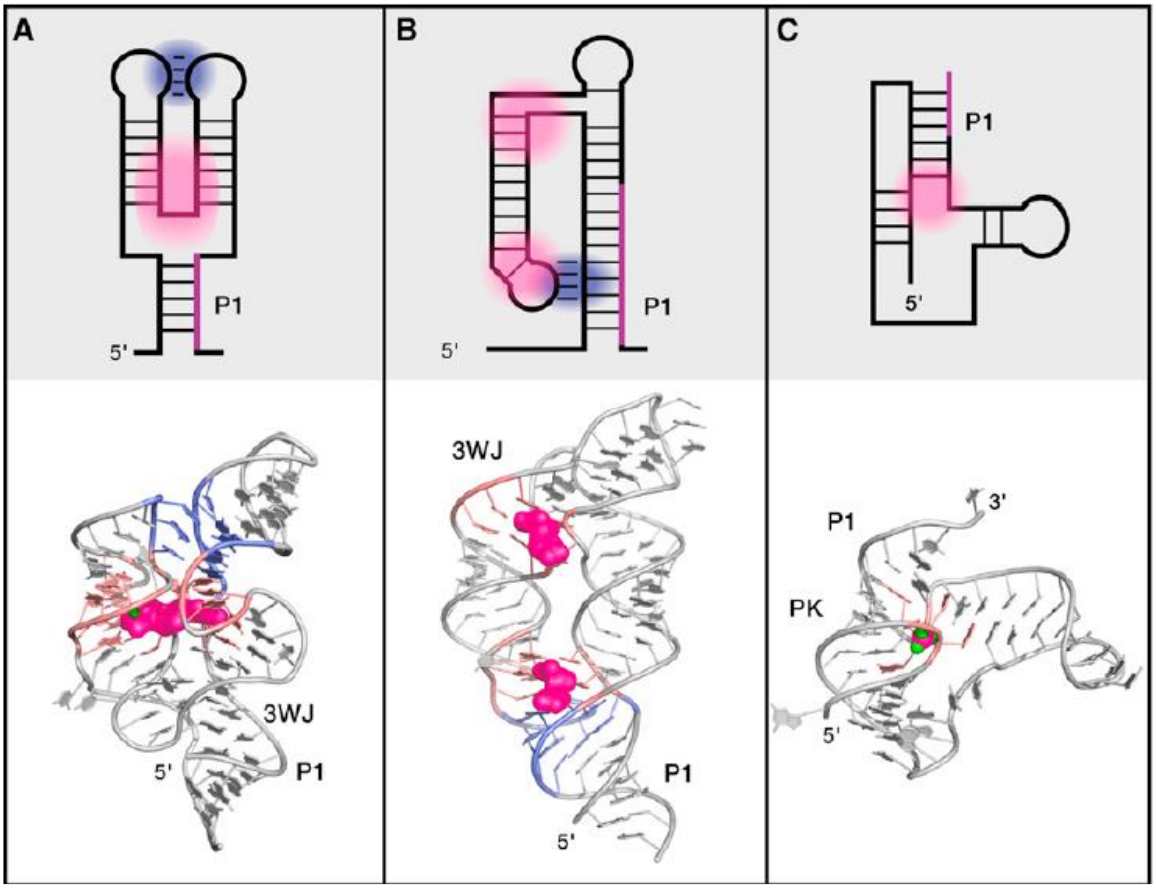

Figure 1.2: Schematic bar (upper panels) and structural (lower panels) representations of three common junctional folds in riboswitches: "straight" (A) with thiamine pyrophosphate, "inverse" (B) with tetrahydrofolate and "pseudoknot" (C) with a fluoride ion. Magenta-colored P1 regions are capable of alternate base-pairing; 3WJ is a three-way junction; pink and blue-shaded areas represent ligand-binding sites and long-distance tertiary interactions, respectively. Adopted from reference 13 with permission.

1.3.2 siRNA and RNA interference

Gene expression can also be regulated by small interfering RNA, or siRNA, which is another type of RNA. These are complementary double-stranded RNA sequences of ~20 base pairs with two-nucleotide overhangs at 3′-OH ends and phosphorylated 5′-termini. siRNA production is catalyzed by the Dicer enzyme from a longer double-stranded construct which is expressed inside the cell or can be introduced there (e.g., by transfection).[20,21] One of the strands, referred to as the guide strand, can base-pair with a complementary region of mRNA and lead to its degradation in a process referred to as RNA interference, or RNAi, which follows a sequence of events. First, to separate the two strands of siRNA from each other, the duplex is bound to the RISC (RNA-Induced Silencing Complex) protein complex which promotes its unwinding while endonucleases degrade the unbound siRNA strand. The guide strand thus becomes free to base-pair to its target mRNA inside the cell. Upon base-pairing, the RISC complex catalyzes ATP-dependent cleavage of mRNA, preventing expression of the protein encoded by that sequence, i.e. silencing a particular gene.[22] Because siRNA targeting is based solely on its complementarity to an mRNA sequence, theoretically any gene can be knocked down through this approach as long as the corresponding mRNA sequence is known. Discovery of RNA interference holds a lot of promise for science and technology and was recognized with the 2006 Nobel Prize in Physiology or Medicine only eight years after publication of results which unequivocally tied the gene silencing observation to introduction of double-stranded RNA into cells.[23] RNAi technology is a “hot” topic today in academia and industry alike: antiviral and cancer treatments, food and agricultural technology

applications as well as numerous other medical and biotechnology uses for RNAi are being developed in the laboratories around the world.[24]

1.3.3 Ribozymes

Ever since the catalytic role of proteins was discovered and the central dogma of molecular biology was formulated, both nucleic acid polymers, DNA and RNA, were considered nature's tools exclusively for storage and transmittal of genetic information or, more specifically, as templates that encode proteins.[25] This principle was put into question by observations of complex secondary and tertiary structure formation by RNA and dismissed after the discoveries of self-cleaving introns and of tRNA-forming protein-RNA complexes in which the RNA component (in some cases with only structural assistance from proteins) was responsible for the catalytic steps. Just as the case with protein enzymes, ribozymes require rigorous geometrical positioning of the catalytic site with respect to the substrate of the reaction. For example, one of the first ribozymes identified was a self-splicing Group I intron from *Tetrahymena thermophila* which sequentially acts on two phosphodiester bonds via S_N2 transesterification mechanisms.[26] In the first of two reactions, a specifically bound and required exogenous guanosine substrate is used to free the 5′ exon, leaving a hydroxyl group at its 3′-terminus (Fig. 1.3). This 3′-OH then serves as the nucleophile to attack precisely at the phosphodiester bond separating the intron sequence from the 3′ exon of mRNA, rejoining the latter to the 5′ exon and completing the splicing process. This fine-tuned excision mechanism requires perfect positioning of the substrates in the catalytic site which in turn must be accurately folded. Just like protein

enzymes, ribozymes adopt stable and well-defined secondary and tertiary structures that can be elucidated by X-ray crystallography.[27]

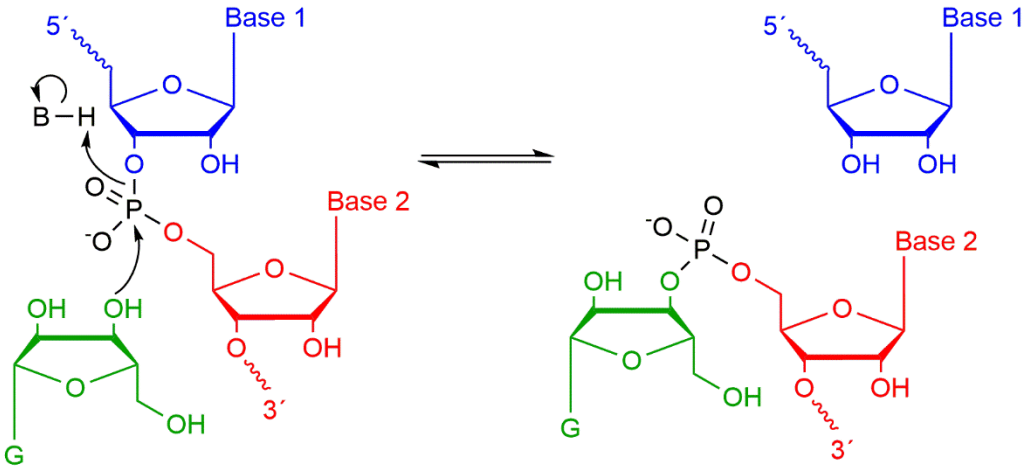

Figure 1.3: Catalytic mechanism of Group I intron for the first of two sequential reactions. Exogenous guanosine is shown in green, last nucleotide of the 5' exon is blue and the first nucleotide of the intron is red.

For his discovery and description of the self-splicing ribozyme, Dr. Thomas Cech was awarded the 1989 Nobel Prize in Chemistry. He shared this award with Dr. Sidney Altman who at around the same time found another catalytic RNA sequence that is an essential part of the RNAse P complex responsible for removing a 5′-leader region from pre-tRNA.[28] The ribozyme discovered by Dr. Altman became one of few examples of natural RNA catalysts capable of multiple turnover, in addition to the Group I intron and the ribosome. RNAse P catalyzes phosphodiester bond hydrolysis to produce mature tRNA molecules and, with some sequence variations, may be found in all living organisms capable of the protein translation process for which tRNA is required.

As mentioned in section 1.2.2, rRNA makes up the ribosome, the most widely known example of a natural ribozyme, which is the sole catalyst of peptide bond formation.[29] The sequence of rRNA in the active site is conserved in all eukaryotes,

indicating the utmost importance of the translation process to life.[30] Briefly, during translation, an aminoacyl-tRNA molecule carrying its cognate amino acid bound to the 3′ terminal hydroxyl enters the A site of the ribosome. This amino acid is added to the growing polypeptide chain's carboxyl end, forming peptidyl-tRNA (Fig. 1.4). After transferring the elongating protein to the next adapter molecule, this uncharged tRNA exits the ribosome at the E site. Although 56 different proteins (in *E. coli*) are required for the ribosome's structural stability, the entire complicated process of translation is orchestrated by the active site's perfectly-positioned rRNA nucleotides, which have remained conserved through billions of years of evolution on Earth.

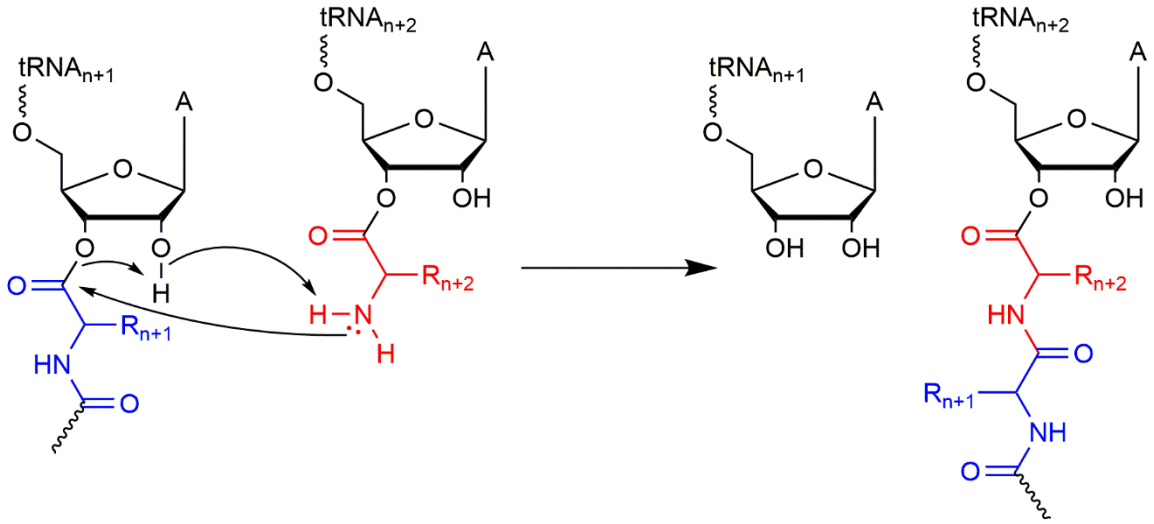

Figure 1.4: Peptide bond formation catalyzed by rRNA in the P site of the ribosome. The new amino acid (red) becomes attached to the already synthesized polymer (blue).

Lacking protein's versatile functional groups, naturally occurring catalytic RNA molecules under physiological conditions require divalent metal ions (most commonly Mg^{2+}) as cofactors for their activity.[31] Metal ions stabilize catalytically active conformations, activate functional groups and participate in the acid-base chemistry of ribozymes. Since high sensitivity of these catalysts to the concentration and identity of

metal ions was experimentally demonstrated, most ribozymes should be placed in the class of metalloenzymes.

After the discoveries of previously unfathomable RNA functionalities in the recent decades, some members of the scientific community began to suggest that early life on Earth was RNA-based in what is widely known as the "RNA World hypothesis".[32-35] In addition to the various catalytic and binding abilities of RNA polymers, nucleotides and their derivatives such as ATP, NADH, acetyl-CoA and others serve as obligatory cofactors for most essential biological processes. The RNA World hypothesis postulates that proteins as the cell's main catalysts emerged later during the development of life on Earth and may have since replaced many functional RNA molecules. This idea suggested that although not present in living organisms, additional RNA catalysts and binders can be identified in the laboratory.

1.4 Non-natural Functional Nucleic Acids

Natural selection is a gradual and often very slow process by which desirable traits become more common in a population with each new generation. At the very basis of natural selection are genes, the units of hereditary information, composed of nucleotides. The process of replication is not perfect and errors, known as mutations, can be made when genes are copied. While most mutations tend to lower an organism's fitness, some may be advantageous. Importantly, without serendipitous errors during copying of genetic information, there would be no chance to improve traits in a population as its old generations are replaced by the new ones through time. Although fascinating because of

its power to advance fitness, the opportunistic process of natural selection is not suitable for a typical biochemical laboratory that aims to evolve a functional nucleic acid on a reasonable time scale. Instead, scientists prefer to start with large populations in which each member is different from every other because a high number of unique members present increases the chance of finding at least one with the ability to perform the desired function. This method of identifying functional molecules in a laboratory is known as in vitro selection and is generally used for RNA and DNA as it is impossible to be applied to amino acid polymers for several reasons.[17] First, due to higher sequence space coverage, a nucleic acid polymer of a given length (e. g., 40 monomers) can have $4^{40} \approx 10^{24}$ possible unique sequences. The number of unique polypeptides of this length is $20^{40} \approx 10^{52}$. Because it is not feasible to make all possible sequence variants for either polymer, the lower number of natural poly(ribo)nucleotide building blocks means that a greater fraction of nucleic acid polymers than polypeptides can be searched. Importantly, direct length comparison is not accurate. Natural functional proteins (on average) are made up of more monomer building blocks than functional single-stranded nucleic acids. The need for testing longer polymers further reduces sequence coverage of polypeptides. Second, random sequences of single-stranded nucleic acids of appreciable length tend to inherently adopt secondary as well as tertiary structures with intricate folds and interactions, the basis of functionality. Polypeptides, on the other hand, require numerous constant stretches of amino acids to support proper protein folding. Building in these conserved regions and preserving their fidelity is not compatible with the in vitro selection approach.

In addition to the conceptual reasons outlined above, there are practical advantages in using nucleic acids over polypeptides. Production of polynucleotides is easier in terms

of cost and time, relative to bacterial expression of proteins. Single-stranded nucleic acid polymers can be prepared via solid-phase synthesis on automated synthesizers. By using a mixture of the four natural nucleotide derivatives, typically in equal proportions, as many as 10^{16} (~20 nmol) unique random sequences can be synthesized in the same vessel and purified together. The total length of each individual sequence is exactly the same while masses differ only slightly. The entirety of all members of this sequence library is referred to as the pool. Flanking the random sequence regions, two constant nucleic acid sections of known identity (each typically 15-20 nucleotides in length) must be present in each member of the pool for amplification purposes. After separating them from non-functional sequences, polymers with desired properties must be amplified in order to increase the total number of each of these unique sequences via polymerase chain reaction (PCR), a technique without which the in vitro selection process would not be viable. In PCR, a complementary sequence binds to the 3′ constant region of each pool member present to serve as a template for a thermostable polymerase enzyme that performs the synthesis of the exact complement of the random region as well as the 5′ conserved region.[36] In the next cycle of PCR, the second template sequence binds to the 3′ end of the reverse complement followed by the polymerase-catalyzed synthesis of the exact copy of the functional nucleic acid polymer. Repetition of these steps during multiple cycles of PCR provides exponential amplification of desired sequences. In addition to their role in template binding for PCR, conserved regions of the pool are used as attachment points for various nucleic acid-conjugated substrates via complementary Watson-Crick base-pairing to each member of the sequence library. Separation of functional sequences from the non-functional ones will be further discussed in Chapter 2, but it is generally performed using polyacrylamide gel

electrophoresis for catalyst selections and via specific binding to molecular targets on a solid support for aptamer identification.

1.4.1 RNA and DNA aptamers

Nucleic acid aptamers are polymers of defined sequences that bind to their specific target molecules. As discussed earlier, riboswitches have aptamer segments which were evolved naturally. Interestingly, the first known aptamers were not found in nature. They were identified prior to the discovery of riboswitches by two separate laboratories via artificial evolution, or in vitro selection, more than a decade earlier.[17,18] More specifically, the process of identifying nucleic acid binders is known as SELEX (Systematic Evolution of Ligands by Exponential Enrichment). In this method, unique single-stranded RNA or DNA sequences with an affinity to a specific molecule are selected out of a random pool by binding their molecular targets appended to a solid support (e.g., an agarose bead).[18] An additional description of aptamer selection process with specific examples is given in Chapter 3 of this thesis.

1.4.2 Artificial ribozymes

After natural catalytic single-stranded RNAs were discovered, scientists became determined to identify unnatural ribozymes. This task necessitated establishing a method of separating a small minority of catalytically active sequences from the rest. For example, in some of the first selections aimed to identify RNA-cleaving and RNA-ligating

ribozymes, the catalytic sequences either acquired or lost a tag, such as a biotin or a thiol functional group, allowing for separating these from the rest.[37] Gel shift also provides separation since the total length of the active sequences changed due to catalytic activity.[38] Because RNA cannot be directly amplified, two additional steps are required for regeneration of the sequence pool. After the separation steps, RNA sequences must be reverse-transcribed into complementary DNA. Following the subsequent PCR step, the now-amplified reverse complement DNA pool must be transcribed into RNA, and this newly regenerated pool is ready to undergo the next round of selection to further enrich the ratio of catalytically active RNA molecules to inactive ones.[39] After sequences with desired activities have emerged from the pool and began dominating it, their identities can be resolved after cloning, testing individual clones for catalytic activity and sequencing the best candidates. Again, due to the presence of an identical conserved region in each molecule of the pool, each clone can be transcribed using a template that is complementary to the constant portion without regard to the composition of the random region. Each transcribed RNA molecule is then presented with a substrate, and the reaction is performed under the conditions used during in vitro selection in a process called screening. DNA clones of RNA sequences capable of catalyzing the reaction of interest above a chosen threshold are then sequenced. The sequences are aligned to find areas of conservation, which may provide hints of the catalytic region's location. Finally, individual ribozymes are prepared for comprehensive studies of catalysis.

The scope of reactions catalyzed by artificial ribozymes is not limited to RNA ligation and cleavage. Catalytic RNA molecules capable of numerous other reactivities have been identified in various laboratories by in vitro selection experiments and this

process continues today.[40-42] Table 1.1 provides several examples of a wide range of reactivities catalyzed by ribozymes identified via the in vitro selection process.

Ribozyme-catalyzed …	Reference	Ribozyme-catalyzed …	Refererence
RNA cleavage	38	Nucleotide synthesis (Fig. 1.5)	43
RNA ligation	44	Amide synthesis (Fig. 1.5)	45
RNA phosphorylation	46	*N*^{7}*G* alkylation	37
RNA capping	47	Aminoacylation	48
RNA branching	49	Michael reaction	50
RNA-protein conjugation	51	Porphyrin metalation	52
Diels-Alder reaction	53	Biphenyl isomerization	54

Table 1.1: Non-comprehensive list of ribozyme-catalyzed reactions and references to reports. Adapted from reference 39.

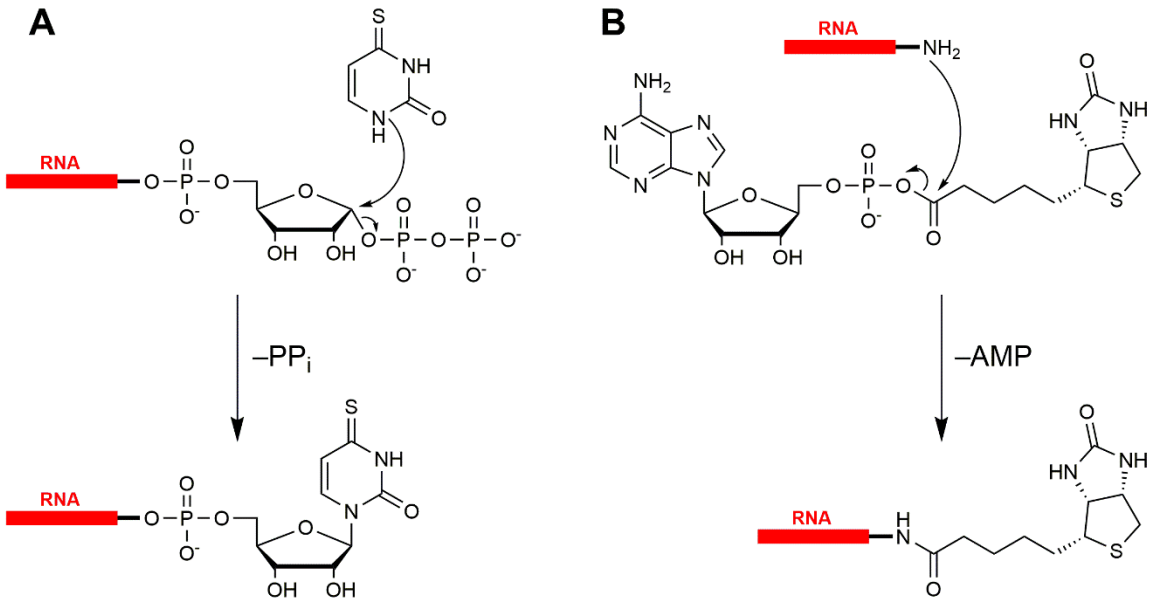

Figure 1.5: Reactions catalyzed by artificial deoxyribozymes: pyrimidine nucleotide synthesis (A) and amide synthesis (B).

1.4.3 Deoxyribozymes

Even though ribozymes both exist in nature and can be identified in the laboratory, deoxyribozymes (also referred to as DNA enzymes or DNA catalysts), single-stranded catalytic DNA molecules, have not been found in living organisms.[55] However, the first

DNA catalyst obtained via in vitro selection was reported in 1994.[56] The two nucleic acid biopolymers are very similar: DNA lacks the 2′-OH and has the thymine nucleobase in place of uracil (a difference of a single methyl group). The use of DNA has additional advantages over RNA:

i) Greater stability since self-cleavage is not possible without 2′-OH of RNA;

ii) DNA is directly amplifiable without the need for transcription and reverse transcription steps during each round of in vitro selection;

iii) Simplicity and cost advantages of DNA solid-phase synthesis versus RNA.

To find deoxyribozymes, scientists employ the same in vitro selection methodology as with ribozymes, but without transcription and reverse transcription steps in each round. The earliest DNA catalyst to be identified, Pb^{2+}-dependent deoxyribozyme, catalyzes the cleavage of the phosphodiester bond at a single ribonucleotide linkage flanked by an all-DNA substrate sequence (Fig. 1.6). For its activity this DNA enzyme strictly requires Pb^{2+} as a divalent metal ion cofactor and displays a rate enhancement of 10^5 over the uncatalyzed background reaction under comparable conditions. Discovery of the Pb^{2+}-dependent DNA catalyst was followed by identification of numerous additional RNA-cleaving deoxyribozymes generally separated into two classes: variants of the 10-23 and 8-27 family of DNA enzymes. The diversity of reactions catalyzed by deoxyribozymes is continuing to expand and a few examples are given in Table 1.2.

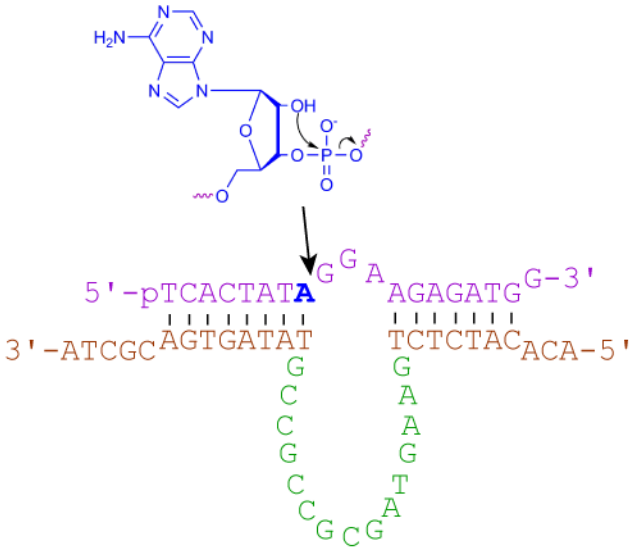

Figure 1.6: Representation of the 10-23 deoxyribozyme complex with its substrate. The substrate contains a single adenosine ribonucleotide (blue) flanked by the all-DNA sequence (teal). Catalytic region of DNA enzyme (green) is flanked by binding arms (brown). Arrow points at the phosphodiester bond cleaved by the 10-23 deoxyribozyme.

Deoxyribozyme-catalyzed …	Reference	Deoxyribozyme-catalyzed …	Reference
RNA cleavage	56	Phosphoramidate cleavage	57
RNA ligation	58	Phosphoramidate linkage	59
RNA branching	60	Oxidative DNA cleavage	61
RNA lariat formation	62	DNA hydrolysis	63
RNA hydrolysis	64	Reductive amination	65
DNA phosphorylation	66	Ester bond hydrolysis	67
DNA adenylylation	68	Aromatic amide bond hydrolysis	67
DNA ligation	69a	Lysine side chain modification	69b
DNA deglycosylation	70	Tyrosine dephosphorylation	71
Thymine dimer photoreveral	72	Tyrosine phosphorylation	73

Table 1.2: Non-comprehensive list of deoxyribozyme-catalyzed reactions and references to reports. Partially adapted from reference 39.

Currently, DNA catalysts for RNA cleavage find the most therapeutic utility potential.[74] This is partially due to the fact that these type of DNA enzymes were some of the first to be identified and partially due to rapid development of RNAi technologies with a similar RNA phosphodiester bond breakage leading to gene silencing. Advantageously, deoxyribozymes do not require additional catalytic functional components such as RISC for siRNA. Additional benefits of using catalytic DNA instead of RNA enzymes include

their greater target specificity and lower costs associated with its synthesis.[75] Unfortunately, the current absence of efficient nucleic acid delivery agents is delaying the therapeutic development of deoxyribozymes as well as other antisense technologies. As a result, no clinical trials of RNA-cleaving deoxyribozymes have been reported yet.[74]

1.5 Protein Side Chain Phosphorylation

Proteins are long functional biological polymers consisting of amino acid building blocks. Most organisms have 20 amino acid monomers encoded by the universal genetic code and connected by peptide bonds. Protein functions are determined by their three-dimensional structures adopted on the basis of amino acid sequence. Structure and function of most proteins can be altered by posttranslational modifications (PTMs), covalent attachment of chemical groups to the side chains of some amino acids.[76] Generally, posttranslational modifications are catalyzed by specific proteins. An example of a PTM is linkage of a phosphate to nucleophilic residues of tyrosine, serine, threonine and histidine where an exogenous ATP or GTP nucleotide serves as the phosphoryl donor.[77] Posttranslational phosphorylation is catalyzed by protein kinases while phosphatases catalyze dephosphorylation. These modifications serve as an important regulatory mechanism by which proteins are either activated or turned off due to resulting structural change. For example, phosphorylation of tyrosine in transmembrane receptor proteins turns on a host of cellular signaling pathways, including those that control translation, transcription, cell cycle progression, and vesicle transport, as well as cytoskeletal and calcium signaling.[78] To study the plethora of cell-regulating interactions, scientists require

innovative approaches to phosphorylate specific amino acid side chains of proteins involved in these events.

1.6 Lanthanide Ions

Collectively known as rare earth elements, the lanthanide series contains fifteen metallic chemical elements from lanthanum (atomic number 57) to lutetium (atomic number 71) and is characterized by filling of the 4f electron shell. Chemical and physical properties of lanthanides are very similar to each other with the most stable oxidation state of 3+, although several of the elements are known to also form stable 2+ and 4+ ions.[79] The main difference arises from filling the f orbital and is referred to as the lanthanide contraction phenomenon, where the radii of ions with the 3+ oxidation state decrease from 103 pm for La^{3+} to 86 pm for Lu^{3+}. All lanthanide ions form coordination complexes with a range of coordination numbers and geometries. Due to their high positive charges, lanthanides act as good Lewis acids and are also known to bind tightly to nucleic acids.[80] Rare earths are becoming a more expensive commodity as they find wider usage in many modern products such as cell phone and hybrid vehicle batteries, magnets, lasers, catalytic converters, catalysts for fuel refining and in glass production as well as in numerous other industrial applications. Due to their unique fluorescent properties, Eu^{3+} and Tb^{3+} ion complexes are used in biological sciences in FRET (fluorescence resonance energy transfer) assays to study interactions on a molecular level.[81] Until recently, lanthanides have not been found to be used in natural biological processes. However, in 2014, Ce^{3+} ions were reported to be essential for energy production by the *Methylacidiphilum*

fumariolicum SolV bacterium.[82] Usually filled by a calcium ion, the active site of the methanol dehydrogenase enzyme of this species is too large for calcium but provides a perfect fit for a cerium ion.

1.7 Research Focus of This Thesis

The research reported in this thesis was aimed at identification of deoxyribozymes with novel cofactor and module requirements. To facilitate the discovery of these nucleic acid enzymes, reactivities already known to be catalyzed by DNA were chosen. Sequence-dependent hydrolysis of phosphodiester bonds in single-stranded DNA by deoxyribozymes[63] and DNA catalysts with tyrosine kinase activity[73] have been reported by other Silverman lab members, and here these reactivities were probed under novel in vitro selection conditions. All new DNA enzymes described in this thesis were identified from unbiased random sequence pools.

In Chapter 2, the ability of Ce^{3+}, Eu^{3+} and Yb^{3+} at low micromolar concentrations to serve as metal ion cofactors in facilitating the discovery of deoxyribozymes was evaluated in the context of DNA cleavage. Attempts at identifying DNA aptamers for small molecules are discussed in Chapter 3. Finally, Chapter 4 is focused on evaluating the utility of incorporating DNA aptamer modules into random sequence pools as binding sites for ATP substrates of tyrosine phosphorylation reactions.

1.8 References

1. Mandelkern, M.; Elias, J. G.; Eden, D.; Crothers, D. M., The dimensions of DNA in solution. *J. Mol. Biol*. **1981**, *152*, 153-161.
2. Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E., Crystal structure analysis of a complete turn of B-DNA. *Nature* **1980**, *287*, 755-758.
3. Kool, E. T., Hydrogen Bonding, Base Stacking, and Steric Effects in DNA Replicaion. *Annu. Rev. Biophys. Biomol. Struct*. **2001**, *30*, 1-22.
4. Crick, F.; Watson, J. D., Molecular Structure of Nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737-738.
5. Mitsui, Y.; Langridge, R.; Shortle, B. E.; Cantor, C. R.; Grant, R. C.; Kodama, M.; Wells, R. D., Physical and enzymatic studies on poly d(I-C)-poly d(I-C), an unusual double-helical DNA. *Nature* **1970**, *228*, 1166-1169.
6. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S., Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.
7. Tennyon, C. N.; Klamut, H. J.; Worton, R. G., The human dystrophin gene requires 16 hours to be transcribed and is cotranscritionally spliced. *Nat. Genet.* **1995**, *9*, 184-190.
8. Jacob. F.; Monod, J., Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **1961**, *3*, 318-356.
9. Lykke-Andersen, S.; Jensen, T. H., Overlapping pathways dictate termination of RNA polymerase II transcription. *Biochimie* **2007**, *89*, 1177-1182.
10. Charette, M.; Gray, M. W., Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **2000**, *49*, 341-351.

11. Pramateftaki, P. V.; Antoniou, P. P.; Typas, M. A., The complete DNA sequence of the nuclear ribosomal RNA gene complex of Verticillium dahliae: intraspecific heterogeneity within the intergenic spacer region. *Fungal Genet. Biol.* **2000**, *29*, 135-143.
12. Poehlsgaard, J.; Douthwaite, S., The bacterial ribosome as a target for antibiotics. Nature reviews. *Microbiology* **2005**, *3*, 870-881.
13. Serganov, A.; Nudler, E., A decade of riboswitches. *Cell* **2013**, *152*, 17-24.
14. Tucker, B. J.; Breaker, R. R., Riboswitches as versatile gene control elements. *Curr. Opin. Struct. Biol.* **2005**, *15*, 342-348.
15. Winker, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R., Control of gene expression by a natral betabolite-responsive ribozyme. *Nature* **2005**, *428*, 281-286.
16. Grundy, F. J.; Henkin, T. M., The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. *Mol. Microbiol.* **1998**, *30*, 737-749.
17. Ellington, A. D.; Szostak, J. W., In vitro selection of RNA molecules that bind secific ligands. *Nature* **1990**, *346*, 818-822.
18. Tuerk, C.; Gold, L., Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA Polymerase. *Science* **1990**, *249*, 505-510.
19. Baker, J. L.; Sudarsan, N.; Weinberg, Z.; Roth, A.; Stockbridge, R. B.; Breaker, R. R., Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* **2012**, *335*, 233-235.

20. Macrae, I. J.; Zhou, K.; Li, F.; Repic, A.; Brooks, A. N.; Cande, W. Z.; Adams, P. D.; Doudna, J. A., Structural basis for double-stranded RNA processing by Dicer. *Science* **2006**, *311*, 195-198.
21. Bernstein, E.; Caudy, A. A.; Hammond, S. M.; Hannon, G. J., Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **2001**, *409*, 363-366.
22. Zamore, P. D.; Tuschl, T.; Sharp, P. A.; Bartel, D. P., RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **2000**, *101*, 25-33.
23. Fire, A.; Xu, S.; A., Kostas, S. A.; Montgomery, M. K.; E., Driver, S. E.; Mello. C. C., Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **1998**, *391*, 806-811.
24. Kupferschmidt, K., A lethal dose of RNA. *Science* **2013**, *341*, 732-733.
25. Crick, F., Central dogma of molecular biology. *Nature* **1970**, *227*, 561-563.
26. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R., Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **1982**, *31*, 147-157.
27. Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A., Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **1996**, *273*, 1678-1685.
28. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S., The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849-857.
29. Cech, T. R., Structural biology. The ribosome is a ribozyme. *Science* **2000**, *289*, 878-879.

30. Ben-Shem, A.; Garreau de Loubresse, N.; Melnikov, S.; Jenner, L.; Yusupova, G.; Yusupov, M., The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* **2011**, *334*, 1524-1529.
31. Johnson-Buck, A. E.; McDowell, S. E.; Walter, N. G., Metal Ions: Supporting Actors in the Playbook of Small Ribozymes. *Metal Ions Life Sci.* **2011**, *9*, 175-196.
32. Gilbert, W., The RNA world. *Nature* **1986**, *319*, 618.
33. Copley, S. D.; Smith, E.; Morowitz, H. J., The origin of the RNA world: co-evolution of genes and metabolism. *Bioorg. Chem.* **2007**, *35*, 430-443.
34. Neveu, M.; Kim, H. J.; Benner, S. A., The "strong" RNA world hypothesis: fifty years old. *Astrobiology* **2013**, *13*, 391-403.
35. Cech, T. R., The RNA worlds in context. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, 1-5.
36. Saiki, R. K.; Gelfand, D. H.; Stoffel, S.; Scharf, S. J.; Higuchi, R.; Horn, G. T.; Mullis, K. B.; Erlich, H. A., Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **1988**, *239*, 487-491.
37. Wilson, C.; Szostak, J. W., In vitro evolution of a self-alkylating ribozyme. *Nature* **1995**, *374*, 777-782.
38. Pan, T.; Uhlenbeck, O. C., A small metalloribozyme with a two-step mechanism. *Nature* **1992**, *358*, 560-563.
39. Silverman, S. K., Artificial functional nucleic acids: aptamers, ribozymes and deoxyribozymes identified by in vitro selection. *Functional Nucleic Acids for Analytical Applications*, Y. Li and Y. Lu, eds.; Springer Science + Business Media, LLC: New York, **2009**, 47-108.

40. Furukawa, K.; Gu, H.; Breaker, R. R., In vitro selection of allosteric ribozymes that sense the bacterial second messenger c-di-GMP. *Methods Mol. Biol.* **2014**, *1111*, 209-220.

41. Moretti, J. E.; Muller, U. F., A ribozyme that triphosphorylates RNA 5′-hydroxyl groups. *Nucleic Acids Res.* **2014**, *42*, 4767-4778.

42. Lau, M. W.; Ferre-D'Amare, A. R., An in vitro evolved glmS ribozyme has the wild-type fold but loses coenzyme dependence. *Nat. Chem. Biol.* **2013**, *9*, 805-810.

43. Chapple, K. E.; Bartel, D. P.; Unrau, P. J., Combinatorial minimization and secondary structure determination of a nucleotide synthase ribozyme. *RNA* **2003**, *9*, 1208-1220.

44. Ekland., E. H.; Szostak, J. W.; Bartel, D. P., Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* **1995**, *269*, 364-370.

45. Wiegland, T. W.; Janssen, R. C.; Eaton, E. B., Selection of RNA amide synthases. *Chem. Biol.* **1997**, *4*, 675-683.

46. Curtis, E. A.; Bartel, D. P., New catalytic structures from an existing ribozyme. *Nat. Struct. Mol. Biol.* **2005**, *12*, 994-1000.

47. Chapman, K. B.; Szostak, J. W., Isolation of a ribozyme with 5′-5′ ligase activity. *Chem. Biol.* **1995**, *2*, 325-333.

48. Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M., Aminoacyl-RNA synthesis catalyzed by an RNA. *Science* **1995**, *267*, 643-647.

49. Tuschl, T.; Sharp, P. A.; Bartel, D. P., Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library. *EMBO J* **1998**, *17*, 2637-2650.

50. Sengle, G.; Eisenführ, A.; Arora, P. S.; Nowick, J. S.; Famulok, M., Novel RNA catalysts for the Michael reaction. *Chem. Biol.* **2001**, *8*, 459-473.

51. Baskerville, S.; Bartel, D. P., A ribozyme that ligates RNA to protein. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9154-9159.

52. Conn, M. M.; Prudent, J. R.; Schultz, P. G., Porphyrin metalation catalyzed by a small RNA molecule. *J. Am. Chem. Soc.* **1996**, *118*, 7012-7013.

53. Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E., RNA-catalysed carbon–carbon bond formation. *Nature* **1997**, *389*, 54-57.

54. Prudent, J. R.; Uno, T.; Schultz, P. G., Expanding the scope of RNA catalysis. *Science* **1994**, *264*, 1924-1927.

55. Breaker, R. R., DNA enzymes. *Nat. Biotechnol.* **1997**, *15*, 427-431.

56. Breaker, R. R.; Joyce, G. F., A DNA enzyme that cleaves RNA. *Chem. Biol.* **1994**, *1*, 223-229.

57. Burmeister, J.; von Kiedrowski, G.; Ellington, A. D., Cofactor-Assisted Self-Cleavage in DNA Libraries with a 3′–5′-Phosphoramidate Bond. *Angew. Chem. Int. Ed.* **1997**, *36*, 1321-1324.

58. Flynn-Charlebois, A.; Wang, Y.; Prior, T. K.; Rashid, I.; Hoadley, K. A.; Coppins, R. L.; Wolf, A. C.; Silverman, S. K., Deoxyribozymes with 2′-5′ RNA ligase activity. *J. Am. Chem. Soc.* **2003**, *125*, 2444-2454.

59. Sachdeva, A.; Silverman, S. K., DNA-catalyzed reactivity of a phosphoramidate functional group and formation of an unusual pyrophosphoramidate linkage. *Org. Biomol. Chem.* **2012**, *10*, 122-125.

60. Wang, Y.; Silverman, S. K., Deoxyribozymes that synthesize branched and lariat RNA. *J. Am. Chem. Soc.* **2003**, *125*, 6880-6881.

61. Carmi, N.; Shultz, L. A.; Breaker, R. R., In vitro selection of self-cleaving DNAs. *Chem. Biol.* **1996**, *3*, 1039-1046.

62. Wang, Y.; Silverman, S. K., Efficient one-step synthesis of biologically related lariat RNAs by a deoxyribozyme. *Angew. Chem. Int. Ed.* **2005**, *44*, 5863-5866.

63. Chandra, M.; Sachdeva, A.; Silverman, S. K., DNA-catalyzed sequence-specific hydrolysis of DNA. *Nat. Chem. Biol*. **2009**, *5*, 718-720.

64. Parker, D. J.; Xiao, Y.; Aguilar, J. M.; Silverman, S. K., DNA catalysis of a normally disfavored RNA hydrolysis reaction. *J. Am. Chem. Soc*. **2013**, *135*, 8472-8475.

65. Wong, O. Y.; Mulcrone, A. E.; Silverman, S. K., DNA-catalyzed reductive amination. *Angew. Chem. Int. Ed.* **2011**, *50*, 11679-11684.

66. Wang, W.; Billen, L. P.; Li, Y., Sequence diversity, metal specificity, and catalytic proficiency of metal-dependent phosphorylating DNA enzymes. *Chem. Biol.* **2002**, *9*, 507-517.

67. Brandsen, B. M.; Hesser, A. R.; Castner, M. A.; Chandra, M.; Silverman, S. K., DNA-catalyzed hydrolysis of esters and aromatic amides. *J. Am. Chem. Soc*. **2013**, *135*, 16014-16017.

68. Li, Y.; Liu, Y.; Breaker, R. R., Capping DNA with DNA. *Biochemistry* **2000**, *39*, 3106-3114.

69. (a) Cuenoud, B.; Szostak, J. W., A DNA metalloenzyme with DNA ligase activity. *Nature* **1995**, *375*, 611-614. (b) Brandsen, B. M., Velez, ; Sachdeva, A.; Ibrahim, N. A.; Silverman S. K., DNA-Catalyzed Lysine Side Chain Modification. *Angew. Chem. Int. Ed.* **2014**, *53*, in press.

70. Sheppard, T. L.; Ordoukhanian, P.; Joyce, G. F., A DNA enzyme with N-glycosylase activity. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7802-7807.

71. Chandrasekar, J.; Silverman, S. K., Catalytic DNA with phosphatase activity. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 5315-5320.

72. Chinnapen, D. J.; Sen, D., A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 65-69.

73. Walsh, S. M.; Sachdeva, A.; Silverman, S. K., DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc*. **2013**, *135*, 14928-14931.

74. Pradeepkumar, P. I.; Höbartner, C., RNA-cleaving DNA enzymes and their potential therapeutic applications as antibacterial and antiviral agents. *From nucleic acid sequences to molecular medicine*, Erdmann, V.; Barciszewski, J., eds.; Springer: Berlin, **2012**; 371-410.

75. Bhindi, R.; Fahmy, R. G.; Lowe, H. C.; Chesterman, C. N.; Dass, C. R.; Cairns, M. J.; Saravolac, E. G.; Sun, L. Q.; Khachigian, L. M., Brothers in arms: DNA enzymes, short interfering RNA, and the emerging wave of small-molecule nucleic acid-based gene-silencing strategies. *Am. J. Pathol*. **2007**, *171*, 1079-1088.

76. Uy, R.; Wold, F., Posttranslational covalent modification of proteins. *Science* **1977**, *198*, 890-896.

77. Burnett, G.; Kennedy, E. P., The enzymatic phosphorylation of proteins. *J. Biol. Chem.* **1954**, *211*, 969-980.

78. Stock, J. B.; Ninfa, A. J.; Stock, A. M., Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* **1989**, *53*, 450-490.

79. Greenwood, N. N.; Earnshaw, A., Cemistry of the Elements (2nd Ed.). Butterworth-Heinemann: Oxform, **1997**; 1233.

80. Franklin, S. J., Lanthanide-mediated DNA hydrolysis. *Curr. Opin. Chem. Biol.* **2001**, *5*, 201-208.

81. Bunzli, J. C.; Piguet, C., Taking advantage of luminescent lanthanide ions. *Chem. Soc. Rev*. **2005**, *34*, 1048-1077.

82. Pol, A.; Barends, T. R. M.; Dietl, A.; Khadem, A. F.; Eygensteyn, J.; Jetten, M. S. M.; Op den Camp, H. J. M., Rare earth metals are essential for methanotrophic life in volcanic mudpots. *Environment. Microbiol.* **2014**, *16*, 255-264.

Chapter 2: Exploration of Requirements of Metal Ions as Cofactors for DNA Catalysts[a]

2.1 Introduction

Deoxyribozymes generally require one or more divalent metal ions as cofactors. Typical metal ion cofactors are Mg^{2+}, Ca^{2+}, Mn^{2+} and Zn^{2+}.[1] In previous work, other members of the laboratory have performed selection experiments with Mn^{2+} together with at least one other divalent metal ion cofactor. Some of these experiments led to successful identification of DNA enzymes that catalyze hydrolysis of phosphodiester bonds in single-stranded DNA.[2-5] For example, 10MD5, the best initial DNA-hydrolyzing deoxyribozyme, requires Zn^{2+} and Mn^{2+} for its catalytic activity and was selected with Mg^{2+} in addition to the two obligatory metal ions. However, via reselection experiments based on partial randomization of the 10MD5 sequence and follow-up studies, a 10MD5 mutant that requires only Zn^{2+} for is activity was identified.[4] Because Mn^{2+} is redox-active and DNA is sensitive to oxidative degradation,[6-9] we wanted to identify new deoxyribozymes that do not require Mn^{2+} for their hydrolytic function. Toward this goal, we decided to evaluate the ability to identify entirely new (starting from a completely random pool) DNA-hydrolyzing deoxyribozymes by selection using Zn^{2+} alone or in combination with Mg^{2+} or Ca^{2+} ions. In a separate experiment, the presence of Mn^{2+} as a single polyvalent metal ion cofactor was also evaluated.

[a] The material described in this chapter has been published and is used here with permission: Dokukin, V.; Silverman S. K., Lanthanide ions as required cofactors for DNA Catalysts. *Chem. Sci.* **2012**, 3, 1707-1714.

In a parallel effort, we evaluated the ability of trivalent lanthanide ions at low micromolar concentrations to support DNA-catalyzed cleavage of DNA. In preliminary work, we attempted to explore the usefulness of these metal ions with existing DNA catalysts. Either Zn^{2+} or Mn^{2+}, initially believed to be required for a new DNA-hydrolyzing deoxyribozyme, was substituted with one of seven different trivalent lanthanide ions in the reaction mixture. In later work for practical reasons, only three lanthanides, Ce^{3+}, Eu^{3+} and Yb^{3+}, were chosen for selection experiments. Due to significant differences in their ionic radii (Ce^{3+} 102 pm, Eu^{3+} 95 pm, Yb^{3+} 86 pm in hexacoordinate environments), these ions could represent early, middle and late regions of the lanthanide row, respectively (Fig. 2.1). The luminescent properties of Eu^{3+} (also Tb^{3+}) are widely known and both lanthanide ions have been used as structural and mechanistic probes for nucleic acids.[10] Identification of new Eu^{3+}-dependent deoxyribozymes could therefore simplify biochemical characterization of catalysts found in these studies or any other future DNA enzymes with remarkable properties selected with and requiring the presence of these luminescent metal ions.

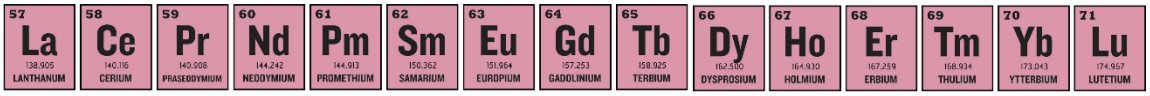

Figure 2.1: Lanthanide row of the periodic table.

2.2 Results and Discussion

2.2.1 Evaluating a known deoxyribozyme's activity with lanthanide ions

Encouraged by tight binding interactions that these metal ions form with nucleic acids and with knowledge that lanthanides have been reported as alternative catalytic cofactors for a particular RNA-cleaving DNA enzyme originally identified to require Pb^{2+},[11] we first decided to test a recently identified deoxyribozyme. 9NL27 is a quintuple mutant of 10MD5 (the first-ever DNA catalyst of single-stranded DNA hydrolysis), reselected from a partially randomized pool for the purpose of expanding the pH range of these deoxyribozymes.[3] Just like 10MD5, the catalysis by 9NL27 was initially observed to require both Zn^{2+} (1 mM optimum) and Mn^{2+} ($K_{d,app}$ ~ 5 mM). A lanthanide ion's ability to substitute for Zn^{2+} or Mn^{2+} (or both ions) in supporting 9NL27 catalysis was initially tested with seven different elements in the 3+ oxidation state (Ce, Eu, Gd, Tb, Dy, Yb and Lu), each at 5 μM, 100 μM and 2 mM. In the case when Mn^{2+} was substituted, all trivalent ions at 2 mM concentrations inhibited the catalysis by 9NL27, while at lower concentrations, DNA hydrolysis activity was observed at a rate comparable to that with both original divalent metal ions. Substitution of a lanthanide cofactor for either Zn^{2+} or both divalent metal ions did not support detectable product formation (Fig. 2.2). This led us to the hypothesis that 20 mM Mn^{2+} can be substituted with a lanthanide ion at a significantly lower concentration. To find the K_d values for lanthanide cofactors, a series of Ce^{3+}, Eu^{3+} and Yb^{3+} concentration ranges were tested. Surprisingly, even at nanomolar concentrations, significant 9NL27 catalysis was observed, raising doubts regarding catalytic roles of lanthanides as cofactors. A negative control experiment with Zn^{2+} alone (without both Mn^{2+} and Ln^{3+}) was performed, revealing the monometallic requirement of

9NL27 catalysis (Fig. 2.2). This observation resulted in significant advances in understanding the characteristics of single-stranded DNA hydrolysis by DNA enzymes by other members of the laboratory.[4]

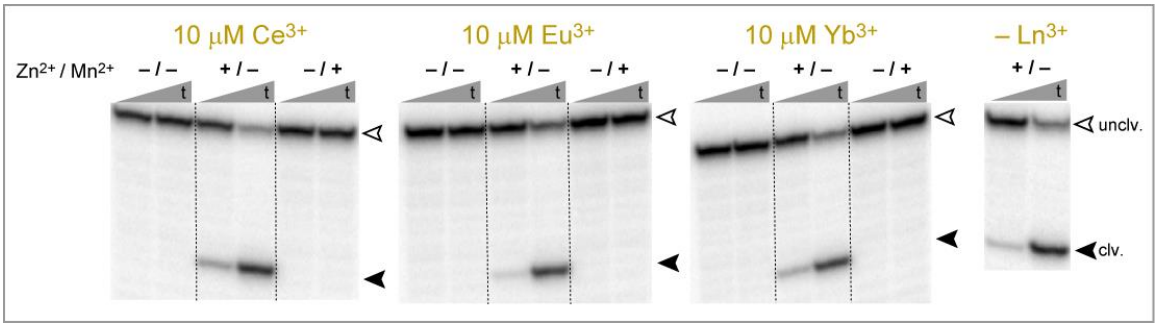

Figure 2.2: PAGE assays of 9NL27 metal ion requirements with 10 μM each Ce^{3+}, Eu^{3+} or Yb^{3+} substituting either one or both of Zn^{2+} (1 mM) and Mn^{2+} (20 mM). Negative control assay with 1 mM Zn^{2+} alone. Timepoints at t = 1 h, 18 h.

2.2.2 Early selection experiments with trivalent lanthanide (Ce^{3+}, Eu^{3+}, Yb^{3+}) and divalent metal ions (Zn^{2+}, Mn^{2+}, Mg^{2+}, Ca^{2+})

Inspired by natural selection, the combinatorial technique of in vitro selection is designed to identify DNA enzymes with novel catalytic properties. Selection of DNA catalysts (e.g., for single-stranded DNA hydrolysis) using this approach involves the following steps (Fig. 2.3). A pool of DNA sequences, each consisting of a random region of the same length (e.g., 40 nucleotides) flanked by constant regions, is generated via solid-phase synthesis. The constant regions are used to bind primers which serve as templates for PCR amplification and for binding of substrate, consisting of primarily complementary DNA sequence and a ribonucleotide GGA segment at the 3′ end. One substrate sequence is then covalently attached to each molecule of DNA pool via reaction catalyzed by the RNA ligase enzyme. In the next step, referred to as the selection step, catalytically active DNA sequences lead to shorter deoxyribozyme-product conjugates that migrate more

quickly on PAGE. Each selection step is performed in the presence of cofactors such as lanthanides as well as other metal ions and a buffer of certain pH. Separation methods depend upon the reactions being catalyzed. In this example, cleavage of the single-stranded DNA substrate anywhere within its central region yields molecules with lower mass and total negative charge, which (due to their faster electrophoretic mobility) are easily separated on PAGE from sequences with uncleaved substrates. The sequences extracted from the gel are then amplified by PCR, followed by ligation and selection steps. These cycles are repeated for several rounds, enriching catalytic sequences iteratively with every round. When catalytic activity of a pool levels off, the individual members of the pool are cloned, sequenced and assayed. Characterization assays are performed intermolecularly, i.e., without covalent attachment of the pool sequences to the substrate (indicated by the orange arrow in Fig. 2.3), in order to find general catalysts because potential DNA cleavage applications would only be useful without the deoxyribozyme-substrate conjugation.

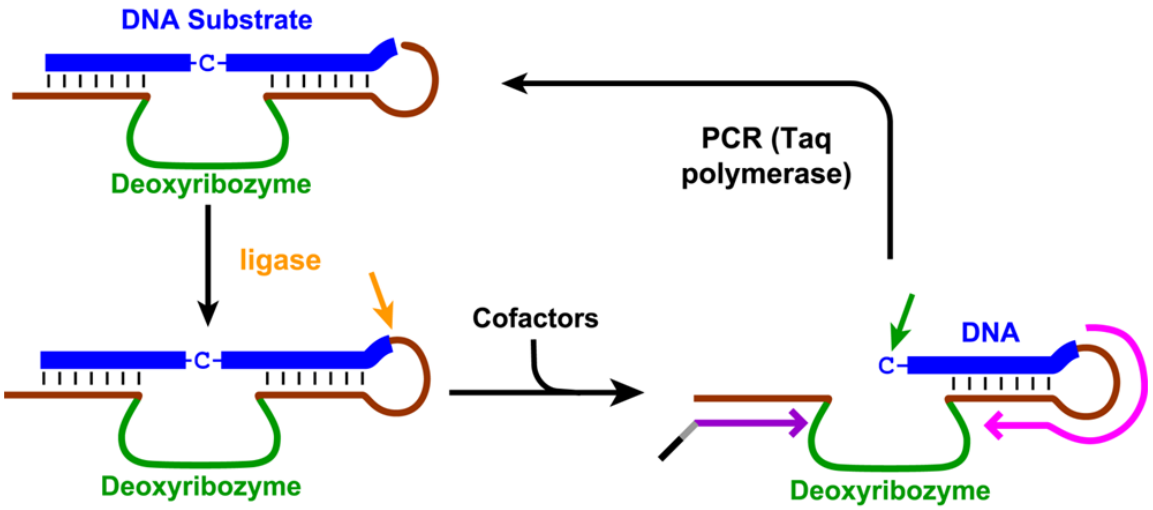

Figure 2.3: In vitro selection strategy to identify deoxyribozymes that cleave DNA. Substrate sequence (blue) with one unpaired C nucleotide is Watson-Crick base-paired to the pool's constant regions (brown) flanking the random region (green). The forward primer (pink) templates synthesis of the pool's sequences during PCR from their reverse complements templated by the reverse primer (purple). The two complementary strands are separated by size due to a polynucleotide tail (black) on the reverse primer separated by a hexaethylene glycol (HEG) spacer (grey) which terminates Taq polymerase-catalyzed synthesis of the pool sequences.

To test viability of lanthanide ions as catalytic cofactors for *de novo* identification of DNA catalysts, a selection experiment, arbitrarily designated SA (according to our laboratory's systematic nomenclature), with a mixture of 1 μM each Ce^{3+}, Eu^{3+} and Yb^{3+} was performed. Because single-stranded DNA cleavage by deoxyribozymes was recently achieved[2] and the methods of selection and further characterization of catalysts already established,[3] this reactivity was chosen for the new study. In addition to the selection with the three lanthanides alone, selection experiments with one of 1 mM Zn^{2+} (SB), 10 mM Mg^{2+} (SC) or 10 mM Ca^{2+} (SD) together with 1 μM lanthanides were performed. For each selection, the key incubation step was performed for 12 h in the presence of either 50 mM (SA, SC, SD) or 70 mM (SB) HEPES, pH 7.5 and 150 mM NaCl at 37°C. Unlike the remainder of the metal ions used here, the 10 × Zn^{2+} stock was prepared in 200 mM

HEPES, pH 7.5 buffer which resulted in the slightly higher HEPES concentration for the SB selection.

Beginning in Round 5, a band corresponding to the cleavage product was observed for the SB selection (Zn^{2+} + Ce^{3+}, Eu^{3+}, Yb^{3+}), corresponding to catalytic activity of 1.5% (Fig. 2.4). The activity continued to increase and reached the value of 36% at Round 7. When the selection step was performed with either Zn^{2+} alone or lanthanides alone, the catalytic activities observed were 25% and 0% (no DNA cleavage), respectively. The remaining selections, SA, SC and SD were iterated further; smearing at the locations of the activity bands appeared in Rounds 7–8 and persisted through Round 11, when these selections were terminated.

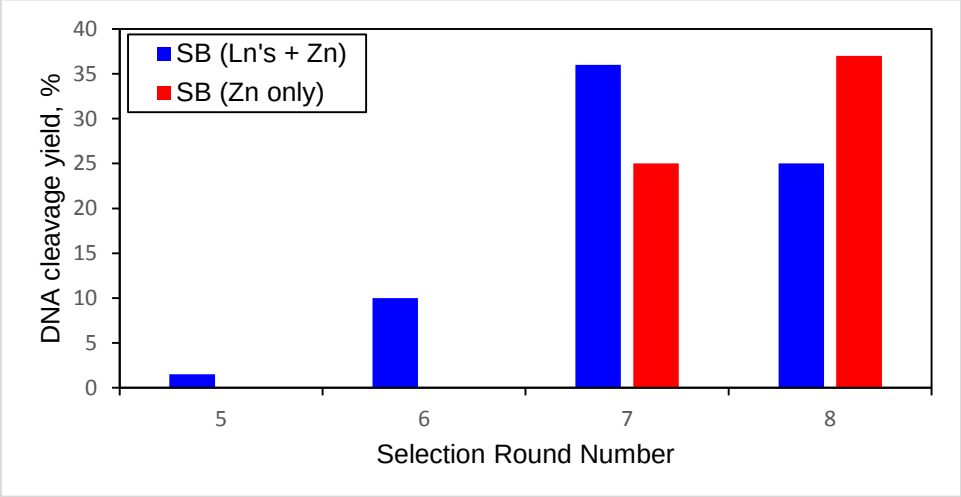

Figure 2.4: Progression of the SB selection experiment.

"Winning" sequences (catalytically active sequences that survived the selection) from the 7SB selection step (round 7, SB selection) were amplified and cloned in *E. coli*. After picking 36 colonies and growing cell cultures, DNA vectors containing inserts of pool sequences were isolated. Sixteen vectors contained single inserts. After PCR

amplification, each sequence was screened by incubation with the single-stranded DNA substrate under selection conditions both with and without the three lanthanides to check whether or not these ions are required for catalysis. Six clones (7SB3, 6, 8, 25, 29 and 32) displayed catalytic activity even without Ce^{3+}, Eu^{3+} and Yb^{3+} while four clones (7SB5, 17, 20 and 23) required lanthanide ions (in addition to Zn^{2+}) to facilitate DNA cleavage. After sequencing all of these clones, the lower-yielding 7SB20 sequence was found to be a single mutant of a more active 7SB5 clone (Fig. 2.5). The other three unique sequences of Ln^{3+}-dependent clones from the formerly random N_{40} region flanked by substrate-binding arms were then synthesized and assayed to determine their metal ion requirements and DNA cleavage yields. Catalytic activity of 7SB5 was much higher-yielding than that of the other two Ln^{3+}-dependent deoxyribozymes which were set aside. Additional experiments revealed that 7SB5-catalyzed DNA cleavage occurs via a deglycosylation pathway instead of the expected hydrolysis reaction. The experimental details of this study are presented in Section 2.2.4 of this chapter, because the protocols were applicable to additional deoxyribozymes. Lanthanide-independent DNA catalysts requiring only Zn^{2+} for their activity that emerged from the SB selection were not studied further due to the indirect way of identifying these DNA enzymes in the presence of lanthanide ions.

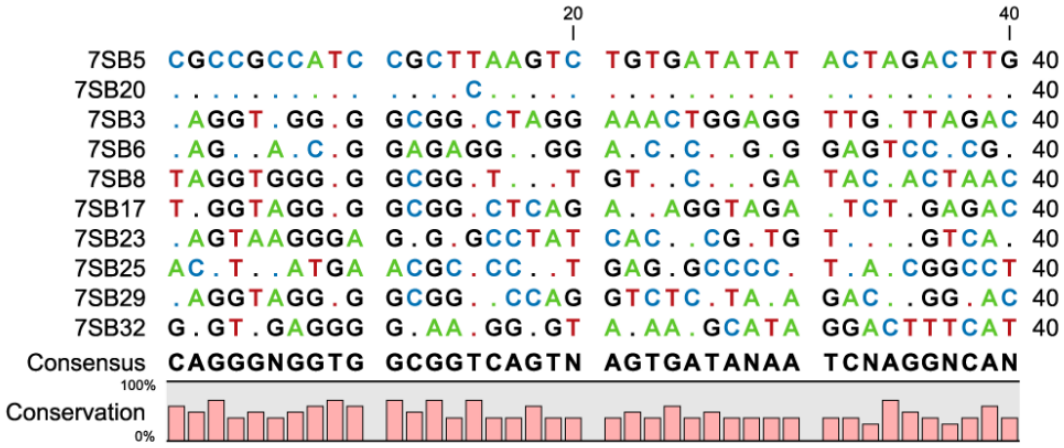

Figure 2.5: Sequence alignment of catalytically active sequences that emerged from the SB selection. All residues identical to those of 7SB5 in the corresponding position are depicted as dots.

Exploration of the ability to identify DNA-hydrolyzing deoxyribozymes without the redox-active Mn^{2+} was pursued further. A total of six selections, designated SL, SM, SN, SP, SQ and SR, were initiated. During the key selection steps, metal ions present during the 14 h incubations were as follows: Mg^{2+}, Mn^{2+} and Zn^{2+} (SL), Mg^{2+} and Zn^{2+} (SM), Ca^{2+} and Zn^{2+} (SN), Mg^{2+} alone (SP), Ca^{2+} alone (SQ) and Mn^{2+} alone (SR); in all cases, the following metal ion concentrations were used: 40 mM Mg^{2+}, 40 mM Ca^{2+}, 20 mM Mn^{2+} and 1 mM Zn^{2+}. All selection reactions were performed in the presence of 150 mM NaCl and either 50 mM CHES, pH 9.0 buffer (SP, SQ selections) or 70 mM HEPES, pH 7.5 buffer (SL, SM, SN, SR selections). The SL selection represented an internal positive control because it was performed under the conditions identical to those under which numerous DNA-hydrolyzing deoxyribozymes were identified. After four selection rounds, the initial activity, as manifested by appearance of a band aligning with the standard band on PAGE, was observed for SL, SM, SN and SR selections with pool yields between 4.0% and 7.7% (Fig. 2.6). The activity increased with yields above 35% for these four selections at Round 6. In both selections that did not have Zn^{2+} during the key

incubation step, the pools did not show catalytic activity through Round 8 and their progression was terminated. All catalytically active sequence pools except for the positive control (SL) selection were cloned from Round 5 PCR products. When sequencing results for DNA-cleaving clones were obtained, it was discovered that 20 out of the 24 sequences were either single mutants of or identical to one of the 7SB25, 7SB29 or 7SB32 sequences obtained in the previous selection effort that require only Zn^{2+} as the sole polyvalent metal ion cofactor. We postulate that these sequences contaminated the pools of the new selections and were amplified during PCR because both primer binding arms of the pools used for the SL-SR selections and the SB selection were identical.

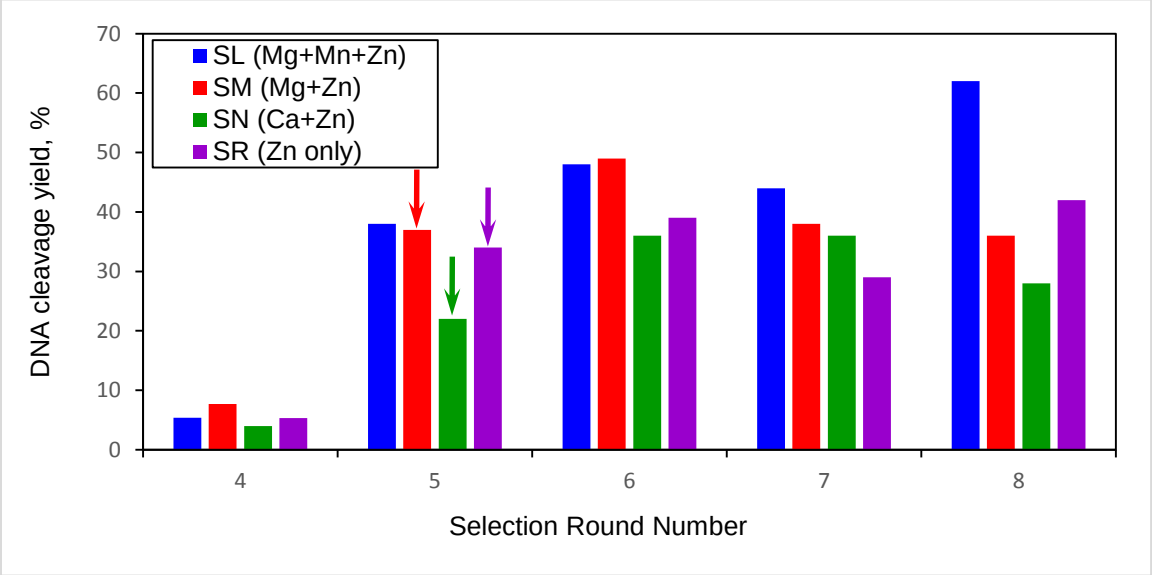

Figure 2.6: Progression of the SL, SM, SN and SR selection experiments. For each selection, the round at which individual deoxyribozymes were cloned is marked with an arrow on the plot.

Because the potential ability of polyvalent metal ions to facilitate deoxyribozyme-catalyzed DNA hydrolysis was still of great interest, in order to avoid a similar contamination problem in the future, a different primer-binding sequence was chosen. The original 15-nucleotides long reverse primer-binding sequence was scrambled (order of

individual nucleotides thoroughly reorganized) to give a new variant. A DNA pool containing the new region and reverse primer complementary to the new binding site were synthesized after verifying (with IDT OligoAnalyzer software) that no significant Watson-Crick base-pairing interactions exist between them and the rest of the constant pool regions as well as the substrate sequence (Fig. 2.7). Because contaminating sequences were propagated during PCR, changing one of the primer regions for the new selections was expected to prevent their amplification. Sequence modifications were not performed on either of the substrate-binding arms, one of which also serves to base-pair with the forward primer during PCR.

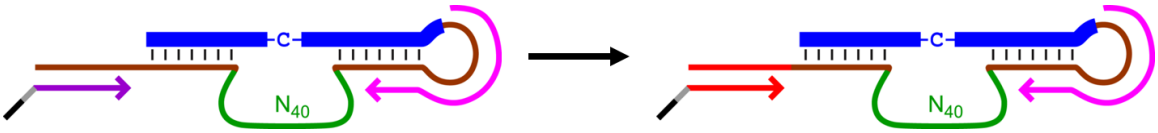

Figure 2.7: An approach to avoid PCR-propagated contamination. Original reverse primer (purple) and its complementary constant region of the pool (brown) sequences were changed (red).

The new pool design was used in the next selection efforts where trivalent (lanthanides) and divalent (Zn^{2+}, Mg^{2+}, Ca^{2+}, Mn^{2+}) metal ions were evaluated for their ability to facilitate DNA hydrolysis by deoxyribozymes. Eight selections were initiated in total: four of these were with 1 mM Zn^{2+} and either one of 10 μM Ce^{3+}, Eu^{3+} and Yb^{3+} or a mix of all three lanthanide ions each at 3 μM concentration. The other selections had either 1 mM Zn^{2+} alone or with one of 40 mM Mg^{2+} or Ca^{2+}; redox-active 20 mM Mn^{2+} as the sole polyvalent metal ion cofactor was also evaluated. Selection reactions were incubated in the presence of either 70 mM (WA-WG) or 50 mM (WH) HEPES, pH 7.5 buffer and 150 mM NaCl. All incubations were performed for 12 h at 37 °C. The bands with appropriate PAGE shifts indicating single-stranded DNA substrate cleavage first

appeared in Round 5 for six out of eight selections and accounted for 3.3% to 8.3% of overall intensity in each lane (Fig. 2.8). Once pool yields increased above a threshold of ~20%, the appropriate pool sequences were cloned. After initial screening and sequencing of catalytically active clones from selections WA and WB, the same contaminating sequences as in the previous selection efforts were discovered. This observation meant that the reverse primer substitution approach was insufficient to prevent propagation of the contaminants which originated in the SB selection. Although lacking the appropriate primer-binding sequence, these Zn^{2+}-dependent DNA catalysts were still copied during PCR: even a single reverse complement could give rise to many copies of the contaminating sequences. The fact that the substrate-binding regions as well as the incubation conditions for all described selection experiments starting with the SB effort were identical means that the contaminating sequences survived the selection en masse, dominated the pools from early rounds and prevented the emergence of novel catalysts.

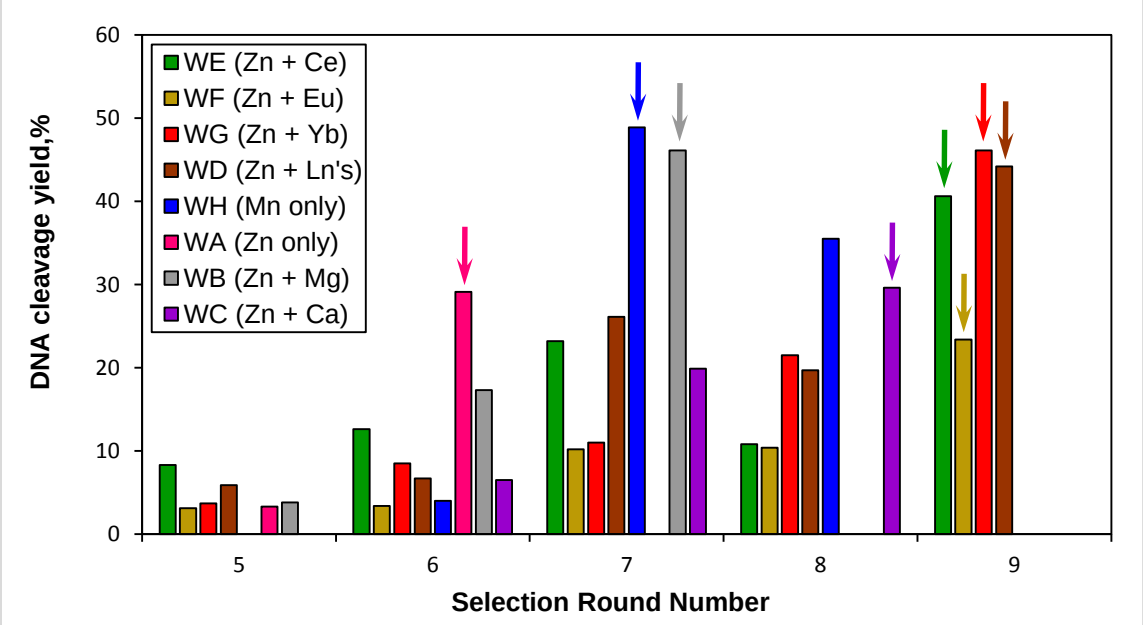

Figure 2.8: Progression of the WA-WH selection experiments. For each selection, the round at which individual deoxyribozymes were cloned is marked with an arrow on the plot.

To prevent the possibility of contamination during the subsequent selection efforts, large portions of both substrate-binding arm sequences of the pool were changed (Fig. 2.9). Only the three nucleotides flanking the N_{40} random region in each arm remained unchanged. A single-stranded DNA substrate sequence complementary to the new binding arms was prepared; the single unpaired C nucleotide remained in the new design. This redesign also necessitated the use of a different forward primer complementary to the new pool's constant region sequence. Using the IDT OligoAnalyzer software, the old substrate binding arms were verified to have no significant base-pairing interaction with the new substrate sequence. With this design, any contaminating sequence that originated in one of the previous selection efforts is prevented from binding the new substrate sequence and is therefore rendered incapable of propagating from one round to the next.

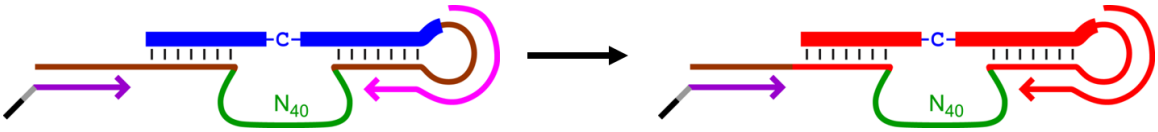

Figure 2.9: An approach to avoid contamination. Original single-stranded DNA substrate (blue) and its complementary constant regions of the pool (brown) as well as forward primer (pink) sequences were changed (red).

Using the redesigned random pools and complementary substrates, eight new selection experiments, designated YE-YM, were initiated. Four of these selections used 1 mM Zn^{2+} and either one of 10 μM Ce^{3+} (YJ), Eu^{3+} (YK) and Yb^{3+} (YL) or a mix of all three lanthanide ions each at 3 μM concentration (YH). The other selections had either 1 mM Zn^{2+} alone (YE) or one of 40 mM Mg^{2+} (YF) or Ca^{2+} (YG); redox-active 20 mM Mn^{2+} as the sole polyvalent metal ion cofactor was also evaluated (YM). Selection reactions were incubated in the presence of either 70 mM (YE-YL) or 50 mM (YM) HEPES, pH 7.5 buffer and 150 mM NaCl. All incubations were performed for 12 h at

37 °C. Each experiment was iterated for numerous selection rounds, leading to substantial DNA cleavage activity by rounds 5 through 7; the progress of selections is presented in Fig. 2.10. Sequences from seven out of eight selection pools were cloned at rounds 6, 7 or 8; due to redundancy, the YH selection (Zn^{2+} + Ce^{3+}, Eu^{3+}, Yb^{3+}), was not studied further. After screening and sequencing of individual clones, no contamination from previous selections efforts was observed. All of the new sequences (Fig. 2.11), therefore, are likely to have arisen from the current selection experiments. Unique sequences from formerly random N_{40} regions flanked by substrate-binding arms were then synthesized and assayed to determine their metal ion requirements, cleavage site specificity and initial rate kinetics.

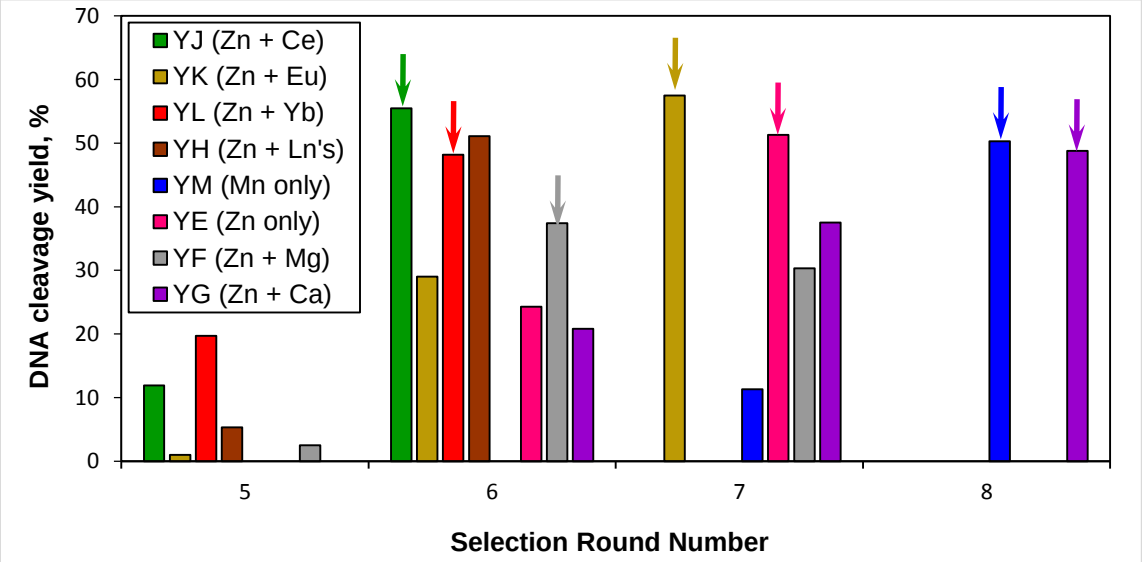

Figure 2.10: Progression of the YE-YM selection experiments. For each selection, the round at which individual deoxyribozymes were cloned is marked with an arrow on the plot.

Figure 2.11: Sequences of the initially random (N_{40}) catalytic regions. Each functional deoxyribozyme comprises the listed catalytic region, surrounded by two Watson-Crick binding arms that interact with the single-stranded DNA substrate.

2.2.3 DNA-catalyzed DNA hydrolysis in the presence of lanthanide ions

For all three selection experiments with Zn^{2+} and one of Ce^{3+}, Eu^{3+} and Yb^{3+}, all of the individual deoxyribozymes were found to require lanthanide ions and to catalyze site-specific DNA hydrolysis. Two deoxyribozymes emerged from the selection with Zn^{2+} + Ce^{3+}: 6YJ10 and 6YJ14 (YJ selection cloned after 6 rounds, clones number 10 and 14). Both DNA catalysts hydrolyzed the phosphodiester bond at the 5′-carbon of the unpaired

C nucleotide, creating 3′-phosphate and 5′-hydroxyl termini, consistent with mass spectrometry results and with the PAGE standard ladder (Fig. 2.12). The same oligonucleotide standards series was loaded on each assay gel. These standards provide a common reference point for comparing gel images and a basis for preliminary assignment of reaction products, albeit with some ambiguity due to near-coincident migration of (i) oligonucleotides of length *n* with 3′-hydroxyl and (ii) oligonucleotides of length *n*+1 with 3′-phosphate (greater mass reduces the migration rate, whereas greater charge increases the migration rate). Additionally, both deoxyribozymes had trace catalytic activity when Ce^{3+} was substituted with Eu^{3+} and no activity detected with Yb^{3+}. Kinetic plots for 6YJ10 and 6YJ14 are presented in Fig. 2.14.

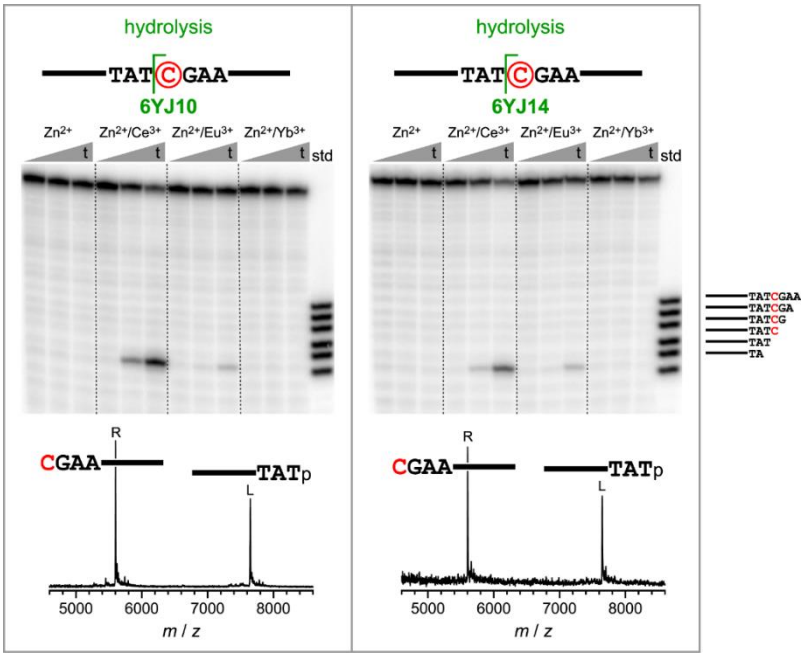

Figure 2.12: Assays of 6YJ10 and 6YJ14 deoxyribozymes that catalyze DNA hydrolysis from the YJ selection experiment that used Zn^{2+} + Ce^{3+} (t = 30 s, 1 h, 16 h). Assays were in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 10 μM $LnCl_3$ and 150 mM NaCl at 37 °C. See Table 2.2 for all mass spectrometry data values. Definitive assignment of reaction products was made on the basis of the mass spectrometry data.

A total of three deoxyribozymes emerged from selection with Zn^{2+} + Eu^{3+}: 7YK24, 7YK34 and 7YK35. Each catalyst hydrolyzed a different phosphodiester bond in the substrate sequence; one DNA enzyme, 7YK34, created 5′-phosphate and 3′-hydroxyl termini on the strands after the hydrolytic cleavage while the other two deoxyribozymes yielded 3′-phosphate and 5′-hydroxyl termini (Fig. 2.13). Cleavage sites are consistent with mass spectrometry results and with the PAGE standard ladder. Both 7YK34 and 7YK35 displayed strong catalytic activity in the presence of either Ce^{3+} or Yb^{3+} in place of the parental Eu^{3+} ion. Interestingly, 7YK24 was more active and had a faster rate when Eu^{3+} was substituted with Yb^{3+}; no detectable hydrolysis was observed with Ce^{3+}. Kinetic plots are presented in Fig. 2.15.

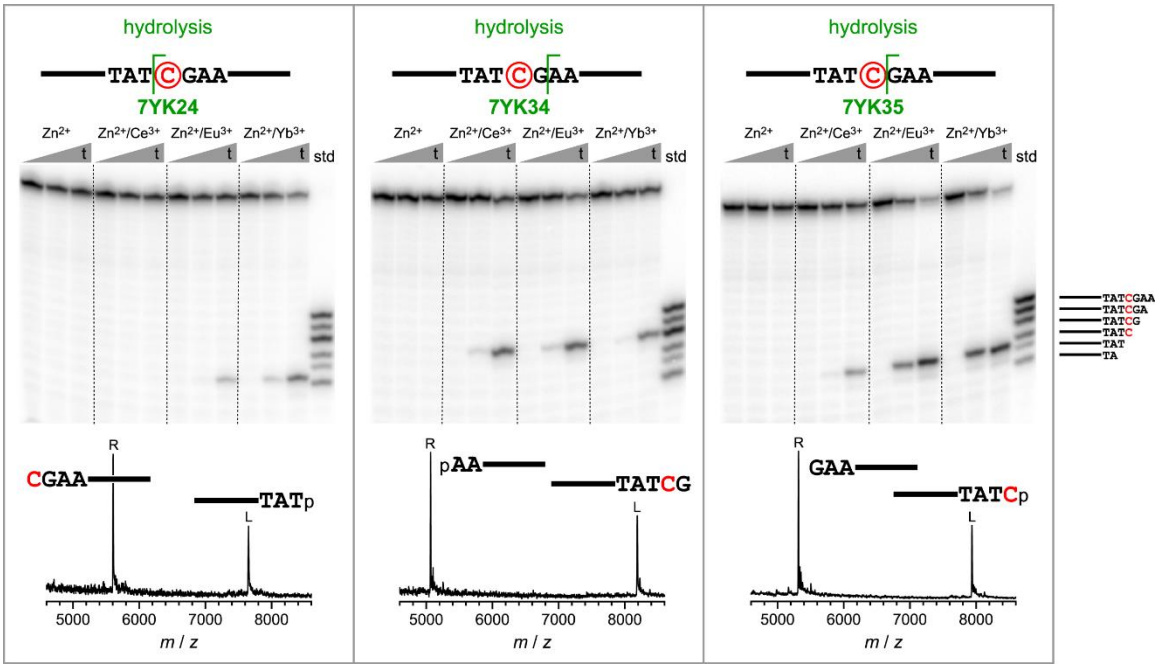

Figure 2.13: Assays of 7YK24, 7YK34 and 7YK35 deoxyribozymes that catalyze DNA hydrolysis from the YK selection experiment that used Zn^{2+} + Eu^{3+} (t = 30 s, 1 h, 16 h). Assays were in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 10 μM $LnCl_3$ and 150 mM NaCl at 37 °C. See Table 2.2 for all mass spectrometry data values. Definitive assignment of reaction products was made on the basis of the mass spectrometry data.

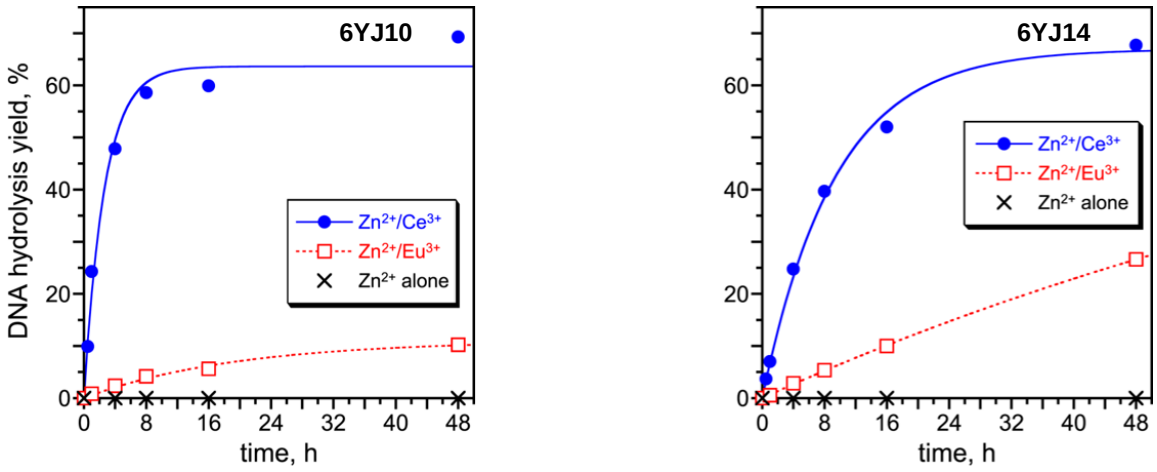

Figure 2.14: Kinetic plots for 6YJ10 and 6YJ14 deoxyribozymes from the YJ selection with Zn^{2+} + Ce^{3+}. k_{obs} values from the first-order curve fits are presented in Table 2.1.

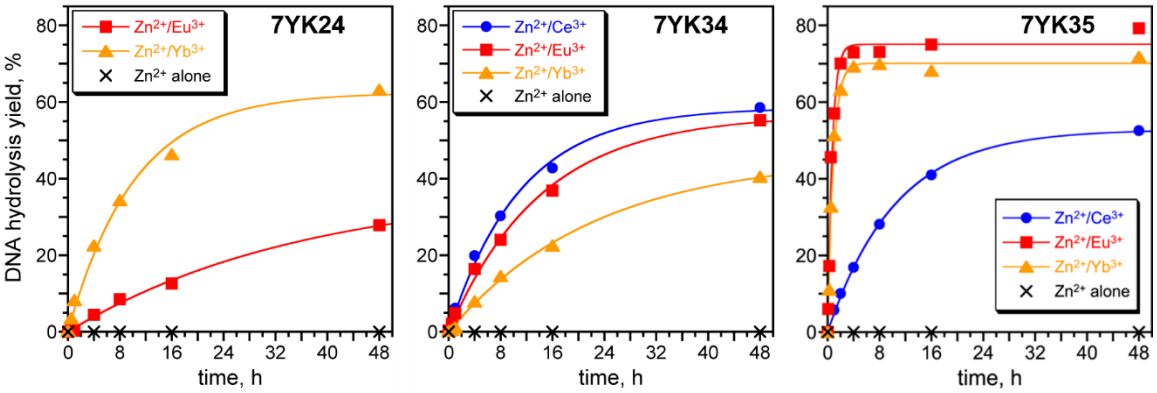

Figure 2.15: Kinetic plots for 7YK24, 7YK34 and 7YK35 deoxyribozymes from the YK selection with Zn^{2+} + Eu^{3+}. k_{obs} values from the first-order curve fits are presented in Table 2.1.

A total of four deoxyribozymes emerged from the selection with Zn^{2+} + Yb^{3+}: 6YL4, 6YL11, 6YL24 and 6YL34. Two different phosphodiester bonds in the substrate sequence were hydrolyzed, each by two different catalysts. All of the 6YL DNA enzymes created 3′-phosphate and 5′-hydroxyl termini after hydrolysis (Fig. 2.15). Cleavage sites are consistent with mass spectrometry results and with the PAGE standard ladder. When evaluated with Zn^{2+}, two out of four deoxyribozymes (6YL4 and 6YL24) had detectable catalytic activity when the parental Yb^{3+} ion was substituted for Eu^{3+}, and no activity was observed with Ce^{3+}.

All nine of the new lanthanide-dependent deoxyribozymes were then separately assayed with the lanthanide ion as the sole polyvalent metal ion cofactor, in the absence of Zn^{2+} despite its presence during selection (50 mM HEPES, pH 7.5, 150 mM NaCl, 37 °C). Both of the Ce^{3+}-dependent deoxyribozymes, all three of the Eu^{3+}-dependent DNA enzymes and two of the four Yb^{3+}-dependent DNA catalysts had no detectable activity without Zn^{2+}, even when 40 mM Mg^{2+} was included to maintain high ionic strength (data not shown). However, two of the Yb^{3+}-dependent deoxyribozymes, 6YL4 and 6YL24, had substantial activity with Yb^{3+} alone in the absence of Zn^{2+}, similar in each case to the activity with Yb^{3+} in the presence of Zn^{2+} (Fig. 2.15). These Yb^{3+}-only activities were maintained in the presence of 40 mM Mg^{2+} (data not shown). Catalysis by both 6YL4 and 6YL24 was achieved by Yb^{3+} alone (no Zn^{2+}) with $K_{d,app}$ on the order of 10 μM, whereas Eu^{3+} was much less effective and Ce^{3+} did not support detectable activity. Kinetic plots are presented in Fig. 2.17.

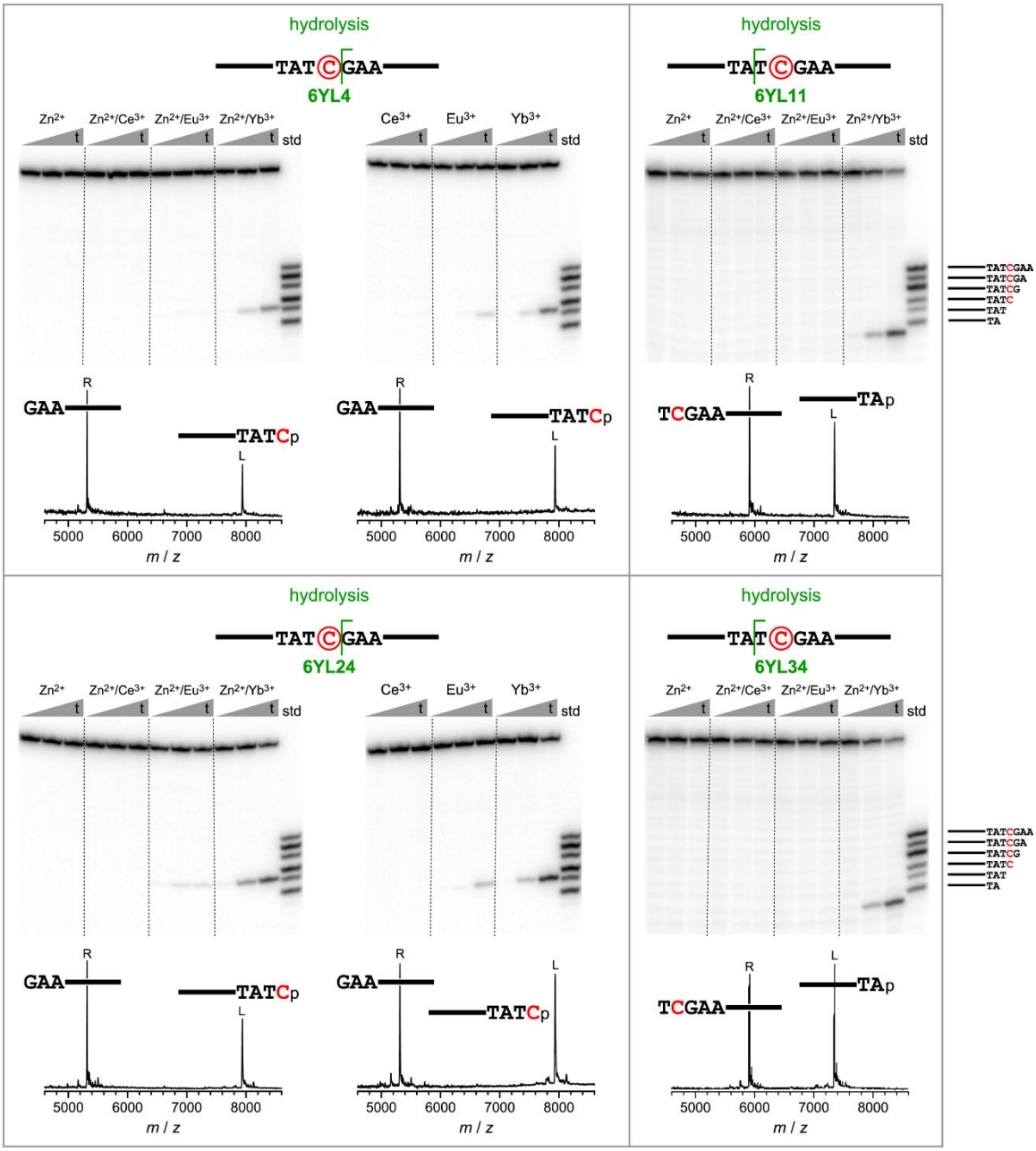

Figure 2.16: Assays of 6YL4, 6YL11, 6YL24 and 6YL34 deoxyribozymes that catalyze DNA hydrolysis from the YL selection experiment that used Zn^{2+} + Yb^{3+} (t = 30 s, 1 h, 16 h). Assays were in 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 10 μM $LnCl_3$ (if included) and 150 mM NaCl at 37 °C. See Table 2.2 for all mass spectrometry data values. Definitive assignment of reaction products was made on the basis of the mass spectrometry data.

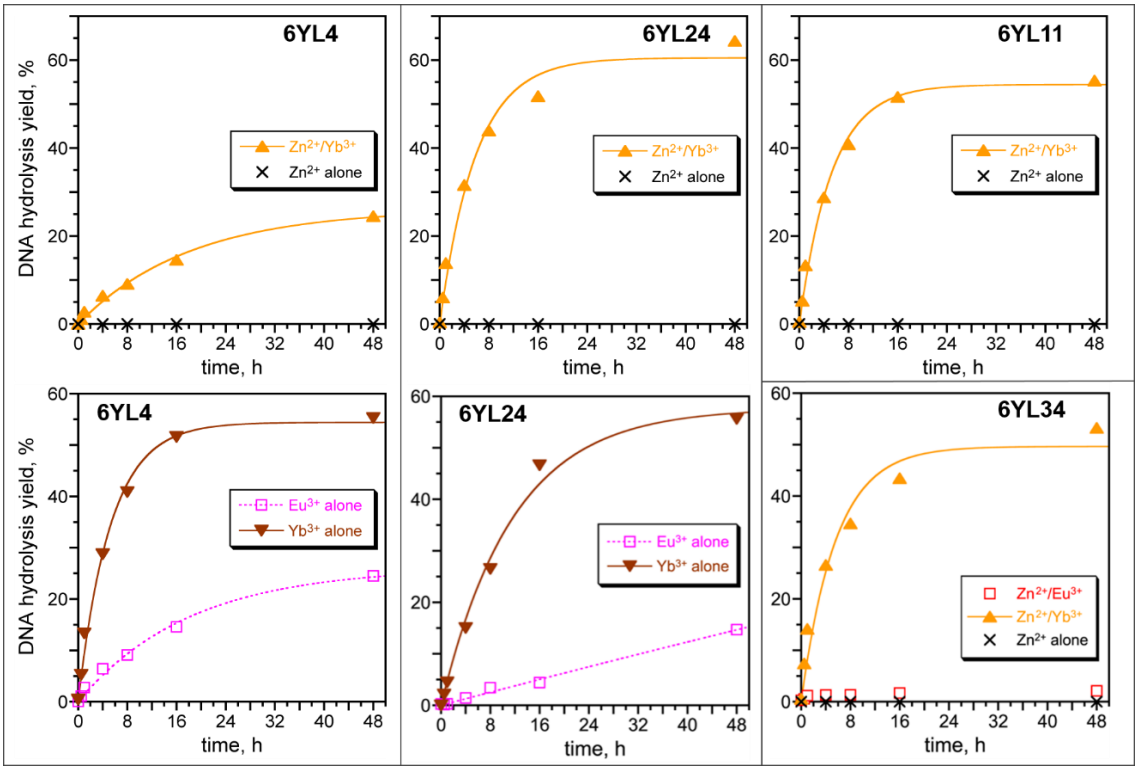

Figure 2.17: Kinetic plots for 6YL4, 6YL11, 6YL24 and 6YL35 deoxyribozymes from the YL selection with Zn^{2+} + Eu^{3+}. k_{obs} values from the first-order curve fits are presented in Table 2.1.

The sharp dependence of many new deoxyribozymes on the lanthanide ion identity is remarkable. We hypothesize that this observation is due to the difference of radius among the three trivalent lanthanide ions (Ce^{3+} > Eu^{3+} > Yb^{3+}). In their tertiary structures, deoxyribozymes identified with a specific lanthanide ion, may have a binding pocket that matches the size of that particular ion. In the case of the Ce^{3+}-dependent 6YJ10 and 6YJ14 deoxyribozymes, substitution of Ce^{3+} with a smaller Eu^{3+} ion significantly inhibits their activity, while they are inactive with an even-smaller Yb^{3+}. We infer that the lanthanide ions with smaller radii cannot coordinate properly in the large binding pockets of 6YJ10 and 6YJ14 to support catalysis (Figs. 2.12 and 2.14). We observe the same trend, but in the opposite direction, for all four DNA catalysts selected with Yb^{3+} (Figs. 2.16 and 2.17).

deoxyribozyme	metal ions	k_{obs}, h^{-1}	yield, %
6YJ10	Zn^{2+}/Ce^{3+}	0.38	69
	Zn^{2+}/Eu^{3+}	0.0060 [a]	10
6YJ14	Zn^{2+}/Ce^{3+}	0.11	67
	Zn^{2+}/Eu^{3+}	0.0072 [a]	27
7YK24	Zn^{2+}/Eu^{3+}	0.028	28
	Zn^{2+}/Yb^{3+}	0.098	63
7YK34	Zn^{2+}/Ce^{3+}	0.091	59
	Zn^{2+}/Eu^{3+}	0.071	55
	Zn^{2+}/Yb^{3+}	0.044	40
7YK35	Zn^{2+}/Ce^{3+}	0.096	53
	Zn^{2+}/Eu^{3+}	1.46	79
	Zn^{2+}/Yb^{3+}	1.15	72
6YL4	Zn^{2+}/Yb^{3+}	0.055	25
6YL11	Zn^{2+}/Yb^{3+}	0.19	55
6YL24	Zn^{2+}/Eu^{3+}	<0.003 [a]	2
	Zn^{2+}/Yb^{3+}	0.18	53
6YL34	Zn^{2+}/Yb^{3+}	0.17	64
6YL4	Eu^{3+}	0.0015 [a]	5
6YL4	Yb^{3+}	0.038 ± 0.009 [b]	48 ± 3 [b]
6YL24	Eu^{3+}	0.0036 [a]	15
6YL24	Yb^{3+}	0.057 ± 0.020 [b]	71 ± 11 [b]

Table 2.1: k_{obs} and yield values for the data plotted in Figs. 2.14, 2.15 and 2.17. k_{obs} values are from first-order kinetic fits unless otherwise noted. Yield values were observed at the 48 h timepoints.

[a] k_{obs} values are from linear fits to the initial time points.

[b] Values represent mean ± sd from four independent experiments.

deoxy-ribozyme [a]	metal ion cofactor(s)	mass L calcd.	mass L found	L error, % (found – calcd.)	mass R calcd.	mass R found	R error, % (found – calcd.)
6YJ10	Zn^{2+}/Ce^{3+}	7648.0	7648.0	0	5603.7	5602.4	–0.02
6YJ14	Zn^{2+}/Ce^{3+}	7648.0	7648.2	+0.003	5603.7	5604.6	+0.02
7YK24	Zn^{2+}/Eu^{3+}	7648.0	7652.2	+0.05	5603.7	5606.4	+0.05
7YK34	Zn^{2+}/Eu^{3+}	8186.4	8188.0	+0.02	5065.3	5066.2	+0.02
7YK35	Zn^{2+}/Eu^{3+}	7937.2	7937.4	+0.003	5314.5	5315.6	+0.02
6YL4	Zn^{2+}/Yb^{3+}	7937.2	7940.2	+0.04	5314.5	5316.1	+0.03
6YL4	Yb^{3+}	7937.2	7934.9	–0.03	5314.5	5312.1	–0.05
6YL11	Zn^{2+}/Yb^{3+}	7343.8	7345.4	+0.02	5907.9	5909.8	+0.03
6YL24	Zn^{2+}/Yb^{3+}	7937.2	7939.0	+0.02	5314.5	5315.7	+0.02
6YL24	Yb^{3+}	7937.2	7936.2	–0.01	5314.5	5313.4	–0.02
6YL34	Zn^{2+}/Yb^{3+}	7343.8	7346.1	+0.03	5907.9	5910.0	+0.04
7YE11	Zn^{2+}	7568.0	7568.4	+0.01	5683.7	5684.3	+0.01
8YG11	Zn^{2+}	8186.4	8189.7	+0.04	5065.3	5067.4	+0.04

Table 2.2: Mass spectrometry assays of cleavage products for DNA-hydrolyzing deoxyribozymes.

[a] See Fig. 2.12 (6YJ deoxyribozymes, Zn^{2+}/Ce^{3+}), Fig. 2.13 (7YK deoxyribozymes, Zn^{2+}/Eu^{3+}), Fig. 2.16 (6YL deoxyribozymes, Zn^{2+}/Yb^{3+} or Yb^{3+} alone) and Fig. 2.23 (7YE11, Zn^{2+}; 8YG11, Zn^{2+}/Ca^{2+}) for each deoxyribozyme's assigned hydrolysis site.

The lanthanide ion concentration dependence of each deoxyribozyme was evaluated with the lanthanide ion used during the selection process along with 1 mM Zn^{2+}, except for the three 7YK DNA catalysts. For 7YK24, the Yb^{3+} concentration dependence was evaluated because in the presence of this lanthanide, the activity of 7YK24 was significantly higher than with Eu^{3+}, the ion with which this deoxyribozyme was selected. Because both 7YK34 and 7YK35 displayed appreciable catalytic activity with all three lanthanide ions, concentration dependence of these DNA catalysts was assayed for each of Ce^{3+}, Eu^{3+} and Yb^{3+}. When 1 mM Zn^{2+} was present, the optimal lanthanide ion concentration for each new deoxyribozyme was ~10 μM. In the absence of 1 mM Zn^{2+}, 6YL4 and 6YL24, which only require lanthanides as polyvalent metal ion cofactors, each displayed a similar level of catalytic activity with 10 and 30 μM Yb^{3+}. Kinetic plots are presented in Figs. 2.18-2.21.

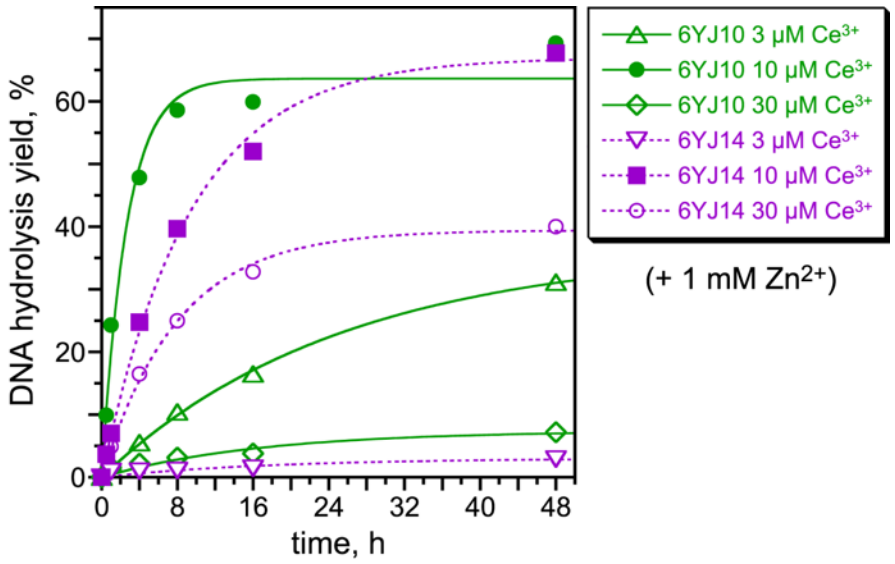

Figure 2.18: Assays of 6YJ10 and 6YJ14 at 1 mM Zn^{2+} and 3, 10, 30 μM Ce^{3+}.

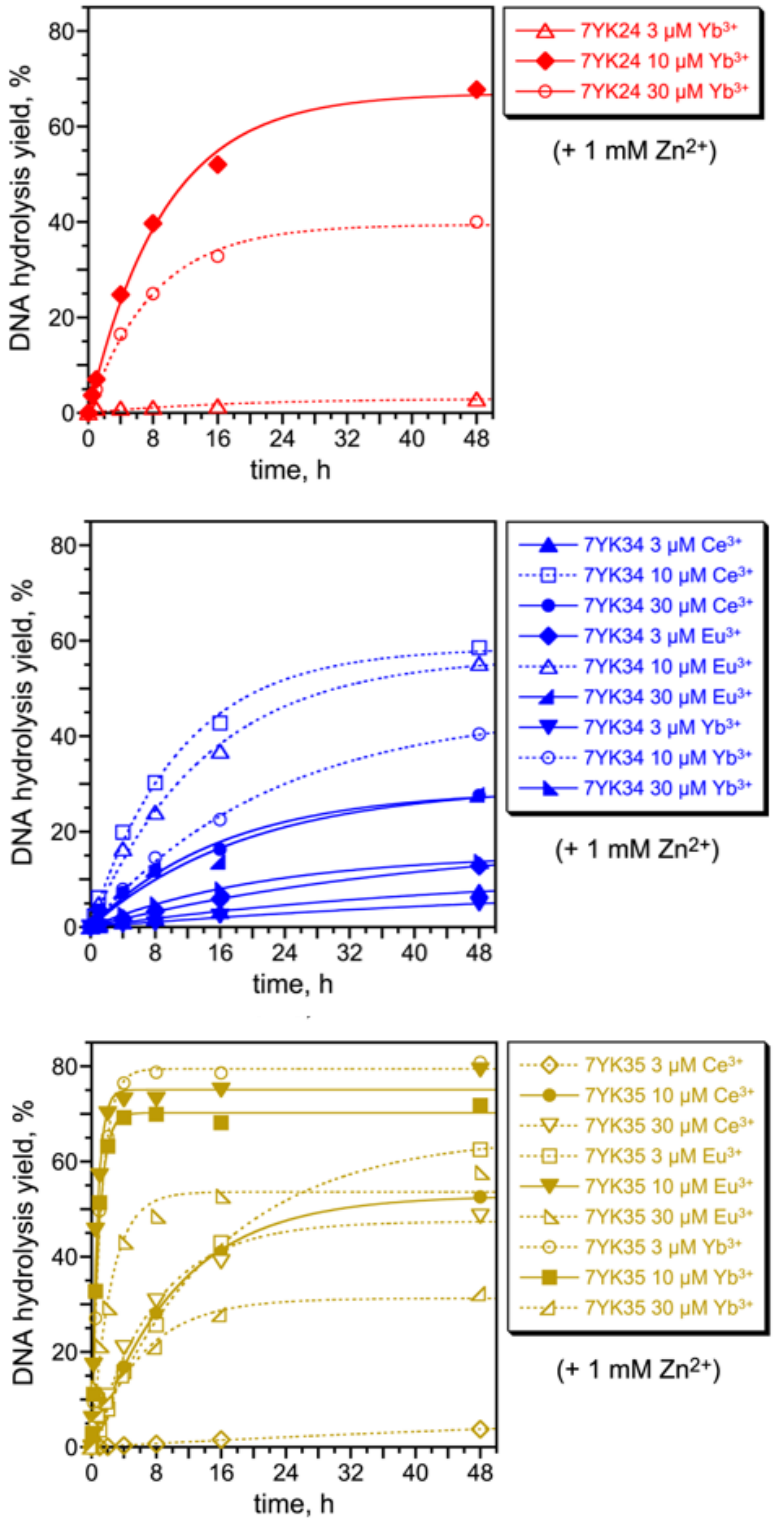

Figure 2.19: Assays of 7YK24, 7YK34 and 7YK35 at 1 mM Zn^{2+} and 3, 10, 30 μM Ce^{3+}, Eu^{3+} or Yb^{3+}.

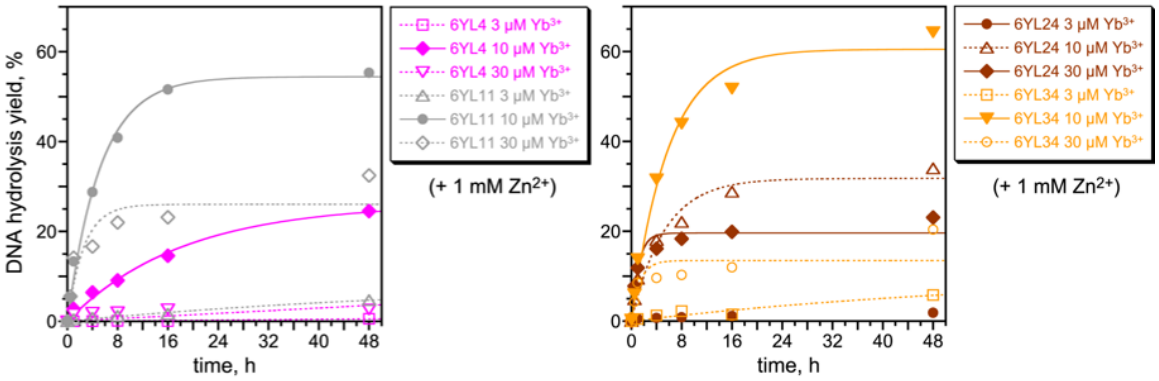

Figure 2.20: Assays of 6YL4, 6YL11, 6YL24, 6YL34 at 1 mM Zn^{2+} and 3, 10, 30 μM Yb^{3+}.

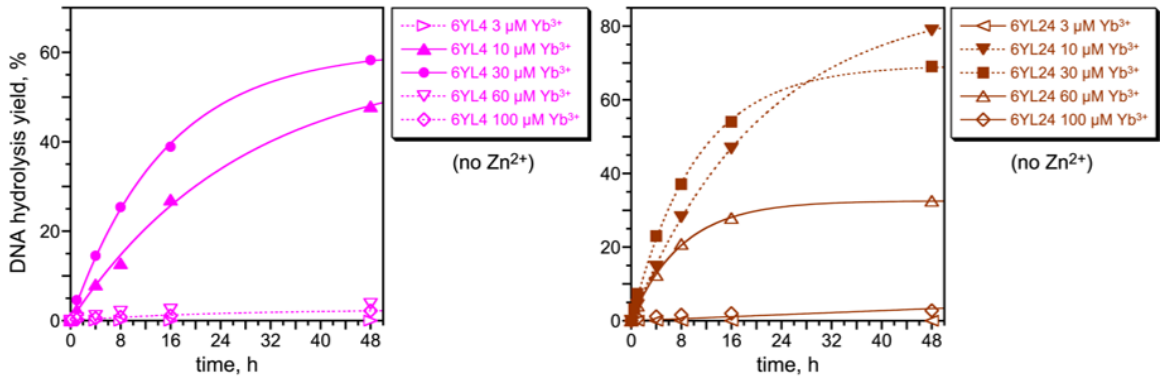

Figure 2.21: Assays of 6YL4 and 6YL24 at 3, 10, 30, 60 or 100 μM Yb^{3+} in the absence of Zn^{2+}.

The dependence of 6YL4 and 6YL24 activity on Yb^{3+} concentration, in the absence of Zn^{2+}, was also quantified by yields at 48 h (Fig. 2.22). The highest yields were achieved at ~10–20 μM Yb^{3+}, whereas steady decrease of activity was observed at higher Yb^{3+} concentrations.

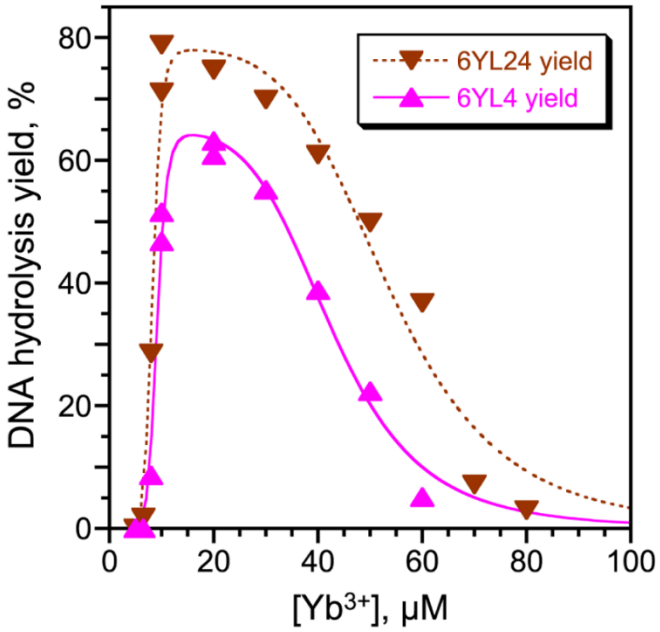

Figure 2.22: Plots of dependence of 6YL4 and 6YL24 catalytic activity on Yb^{3+} concentration as quantified by yields at 48 h.

To create the plots in the above figure, each data set was fit to the following equation:

$$Y = Y_{\max} \cdot \frac{C^n}{K_{\text{activ}}^n + C^n} \cdot \left(1 - \frac{C^m}{K_{\text{inhib}}^m + C^m}\right)$$

where Y = yield, C = Yb^{3+} concentration, K_{activ} and K_{inhib} are apparent binding constants for Yb^{3+} activation and inhibition, respectively, and n and m are Hill coefficients for Yb^{3+} binding and activation, respectively. This equation was derived from a model in which n activating Yb^{3+} ions must bind for the deoxyribozyme to adopt a productive conformation for catalysis, whereas m inhibitory Yb^{3+} ions must not bind for the deoxyribozyme to adopt

an productive conformation and, moreover, binding of the *n* activating and *m* inhibitory Yb^{3+} ions are independent events (and therefore the relevant probabilities multiply). Fit values were as follows: 6YL4 Y_{max} 62%, K_{activ} 9 μM, K_{inhib} 44 μM, *n* 14, *m* 6; 6YL24 Y_{max} 75%, K_{activ} 8 μM, K_{inhib} 55 μM, *n* 14, *m* 6. The precise numerical values are less important than the overall finding that data for both deoxyribozymes are fit well by a simple model in which Yb^{3+} activation is highly cooperative with ~10 μM apparent binding constant and Yb^{3+} inhibition is somewhat less cooperative with several-fold higher apparent binding constant. We cannot discount the possibility that the "inhibition" observed at high Yb^{3+} concentration is instead a nonspecific effect due to Yb^{3+} precipitation or aggregation rather than true inhibition via binding of Yb^{3+} to the deoxyribozyme. However, we have no particular evidence for Yb^{3+} precipitation or aggregation, and others have probed deoxyribozyme activity with lanthanide ions at concentrations considerably higher than we have used here (up to 1.5 mM at pH 7.0).[11]

Because of the prior report of a Pb^{2+}-dependent DNA enzyme that can use lanthanide ions in place of Pb^{2+},[11-18] for each of the new deoxyribozymes we also evaluated replacement of the lanthanide ion with 1 mM Pb^{2+}, either with or without the additional inclusion of Zn^{2+}; in all cases, no activity was observed. Also, because each of Ce^{3+}, Eu^{3+} and Yb^{3+} have accessible 2+ oxidation states,[19] we checked the activities of each of the nine new lanthanide-dependent deoxyribozymes with La^{3+}, Gd^{3+} and Lu^{3+}, each of which is adjacent to one of Ce^{3+}, Eu^{3+}, or Yb^{3+} (respectively) but cannot readily access its 2+ oxidation state. In all cases, substantial DNA hydrolysis activity was retained, indicating that the catalysis does not depend upon accessing the lanthanide 2+ oxidation state.

2.2.4 DNA-catalyzed DNA cleavage in the absence of lanthanide ions

DNA catalysts identified from the four selections without lanthanide metal ions were also characterized. Only two deoxyribozymes were found to catalyze phosphodiester bond hydrolysis: 7YE11 (from the Zn^{2+}-alone selection) and 8YG11 (from the Zn^{2+} + Ca^{2+} selection) (Fig. 2.23). The remaining 18 DNA enzymes catalyze deglycosylation (cleavage of the N-glycosidic bond), eventually leading to single-stranded DNA strand scission after two spontaneous β-elimination reactions. The two reactivities can be discerned by MALDI mass spectrometry because the sets of two cleavage products from each mechanism are not

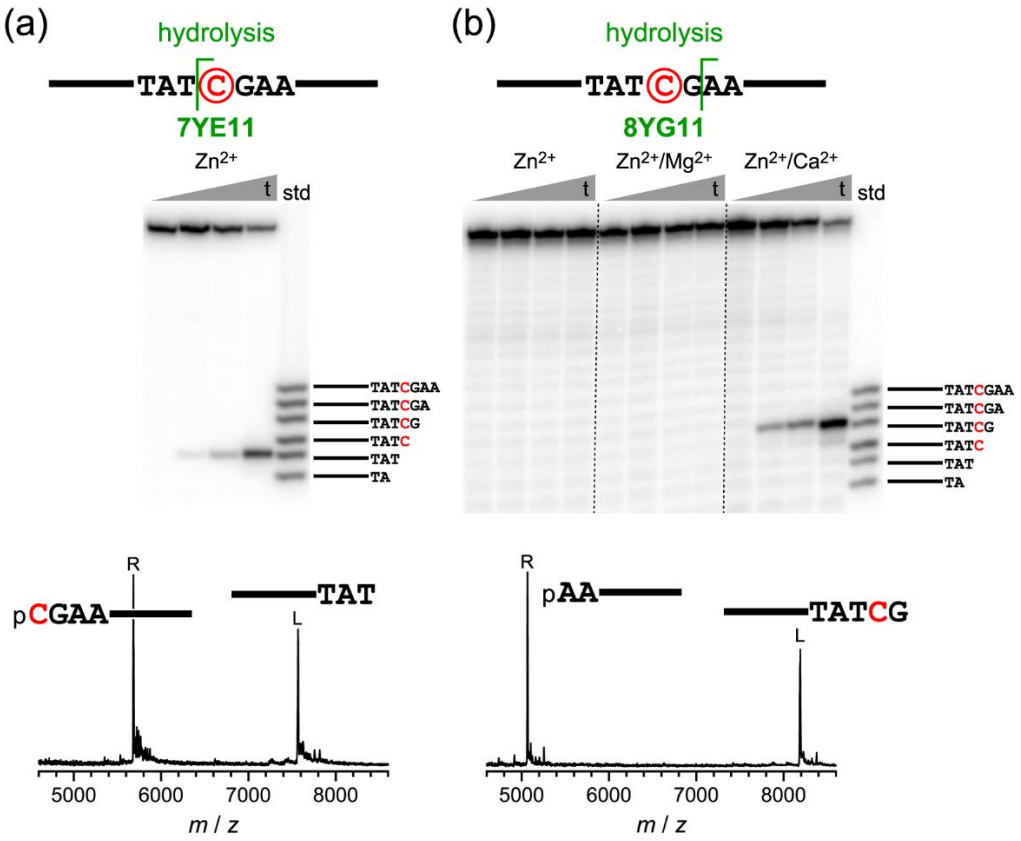

Figure 2.23: Assays of two DNA-hydrolyzing deoxyribozymes identified in the absence of lanthanide ion cofactors. Representative timepoints at t = 30 s, 30 min, 2 h and 16 h under single-turnover conditions. (a) 7YE11 deoxyribozyme, identified with Zn^{2+} alone. (b) 8YG11 deoxyribozyme, identified with Zn^{2+} + Ca^{2+}. Presentation of data is similar to that shown in Figs. 2.12, 2.13 and 2.16.

identical: deglycosylation results in a loss of one nucleobase and one sugar residue while there is no loss of functional groups due to hydrolytic cleavage of a phosphodiester bond (Fig. 2.24).

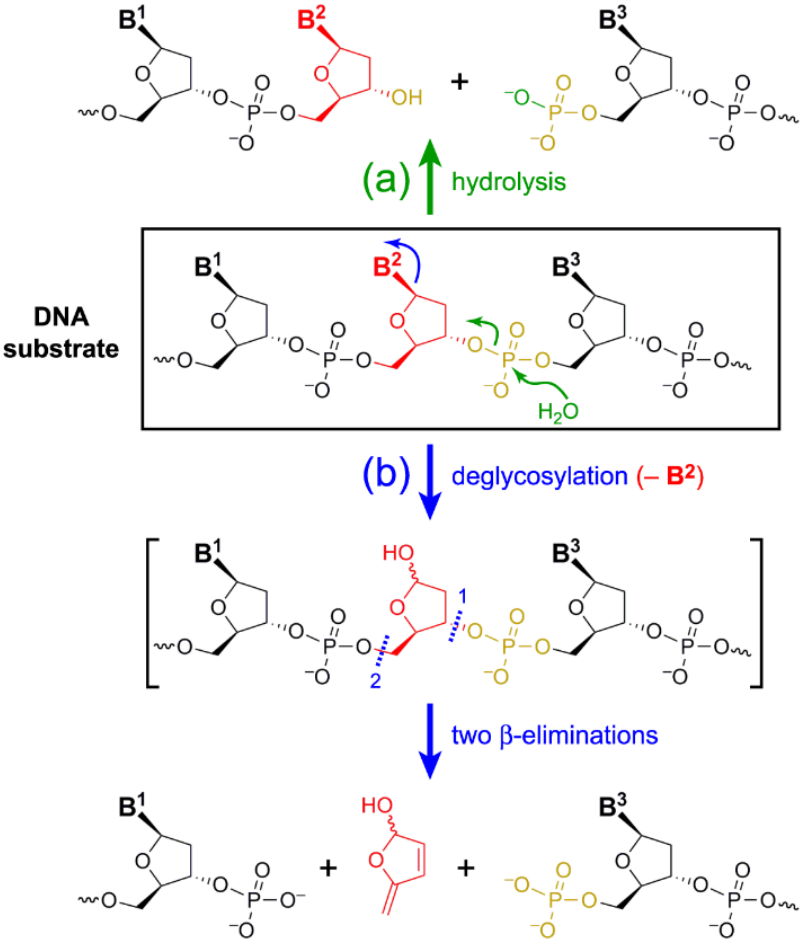

Figure 2.24: Comparison of two pathways for DNA cleavage. (a) Hydrolysis, leading to either 3′-hydroxyl + 5′-phosphate termini (shown) or 3′-phosphate + 5′-hydroxyl termini (not shown). (b) Deglycosylation followed by strand scission via two β-elimination reactions, leading to 3′-phosphate + 5′-phosphate termini and a "missing" nucleoside. B = nucleobase.

All but one (8YM4) of these DNA enzymes catalyze non site-specific deglycosylation of nucleobases at multiple positions within the central region of the substrate sequence, including the unpaired C (Fig. 2.25). 8YM4 site-specifically

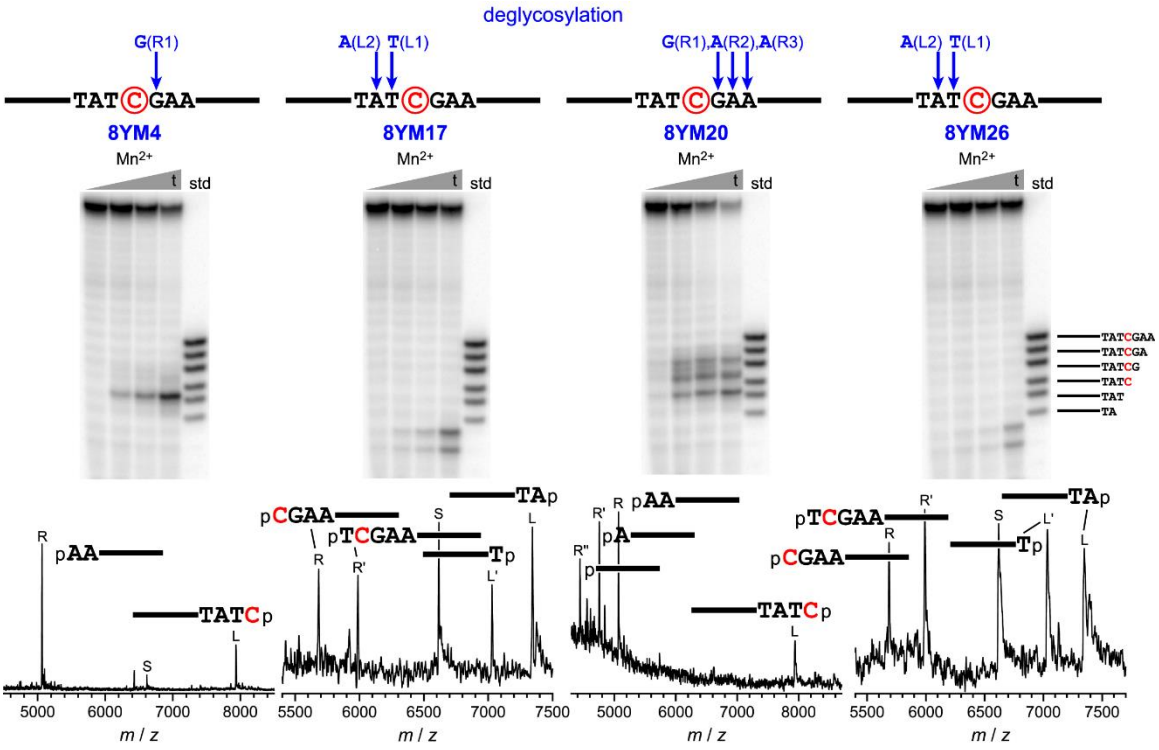

Figure 2.25: Assays of 8YM4, 8YM17, 8YM20 and 8YM26 deoxyribozymes that catalyze DNA cleavage by deglycosylation and β-elimination from the YM selection experiment that used Mn^{2+} alone (t = 30 s, 30 min, 2 h, 16 h). The arrows above each sequence mark the site(s) of deglycosylation, as assigned on the basis of the MALDI mass spectrometry data. Nucleotides are numbered as part of either the L (left) or R (right) portion of the substrate, counting outward. See Table 2.3 for all mass spectrometry data values. Mass spectrometry peaks labeled S correspond to uncleaved substrate (*z* = 2). In all cases, the assigned deglycosylation sites were consistent with the PAGE standard ladders. Definitive assignment of reaction products was made on the basis of the mass spectrometry data.

deglycosylates the G nucleobase directly to the right of the unpaired nucleotide. Most of these DNA catalysts cleave the glycosidic bonds of either one of two adjacent nucleotides; 8YM20 acts on one of three adjacent nucleobases. Many deoxyribozymes with two cleavage sites display a preference for one nucleobase over the other. Observation of depurination and depyrimidination at adjacent nucleotides (e.g., 8YM17 or 8YM26; Fig. 2.26) catalyzed by the same deoxyribozyme suggests a deglycosylation mechanism that is indifferent to the structure of the nucleobase, but more studies are needed to probe these mechanism(s). The finding of depyrimidination (at T or C) rather than depurination

(at A or G) by 8YM17 and 8YM26 (Fig. 2.26) as well as several of the Zn^{2+}/Ca^{2+}-dependent deoxyribozymes (Fig. 2.28) is especially curious. Depyrimidination is generally more difficult to catalyze than is depurination, because pyrimidine nucleobases lack the purine N7 atom that can be protonated or otherwise activated to assist deglycosylation.[20,21] Although the uncatalyzed deglycosylation rates of purine and pyrimidine nucleosides in DNA are similar[22] and uracil DNA glycosylase is well known for damage repair,[23-27] natural enzymatic examples of depyrimidination of cytidine or its analogues are unommon.[21,28,29]

Although loading PAGE standard ladders alongside assay samples afforded a good estimate of the cleavage site location within the substrate, masses of product fragments obtained by MALDI provided truly conclusive evidence. In several cases, the left-hand (L) product formed upon deglycosylation and β-elimination was observed as an L+58 peak, i.e., a peak with mass 58 higher than expected for the L product with a simple 3′-phosphate group. Typically both the L and L+58 peaks were observed. Per Fig. 2.24, the first β-elimination reaction after deglycosylation releases the right-hand (R) oligonucleotide product, which is always observed cleanly by mass spectrometry in its simple 5′-phosphate form. The ensuing, second β-elimination reaction should release L. However, if the excised nucleoside sugar ring (red in Fig. 2.24) is partially fragmented rather than fully released from the L product fragment, then the L product fragment will have slightly higher mass than expected. There is precedent for such partial sugar fragmentation, at least for oxidatively induced DNA cleavage reactions.[30,31] In the present examples, although the available metal ions—either Zn^{2+} alone or Zn^{2+} with Mg^{2+} or Ca^{2+}—are not redox-active,

we speculate that a similar sugar ring fragmentation occurs (via unknown mechanism), leading to the observed L+58 = L+$C_2H_2O_2$ products.

One of the five deoxyribozymes identified from selections with Zn^{2+} alone, 7YE11, catalyzes phosphodiester bond hydrolysis. Three of the other four DNA catalysts (7YE5, 7YE8, 7YE19) preferably facilitate depurination of the G nucleobase immediately adjacent to the unpaired nucleotide while the neighboring A is the secondary deglycosylation site. For 7YE2, the depurination sites are the same, but site preferences are reversed (Fig. 2.26).

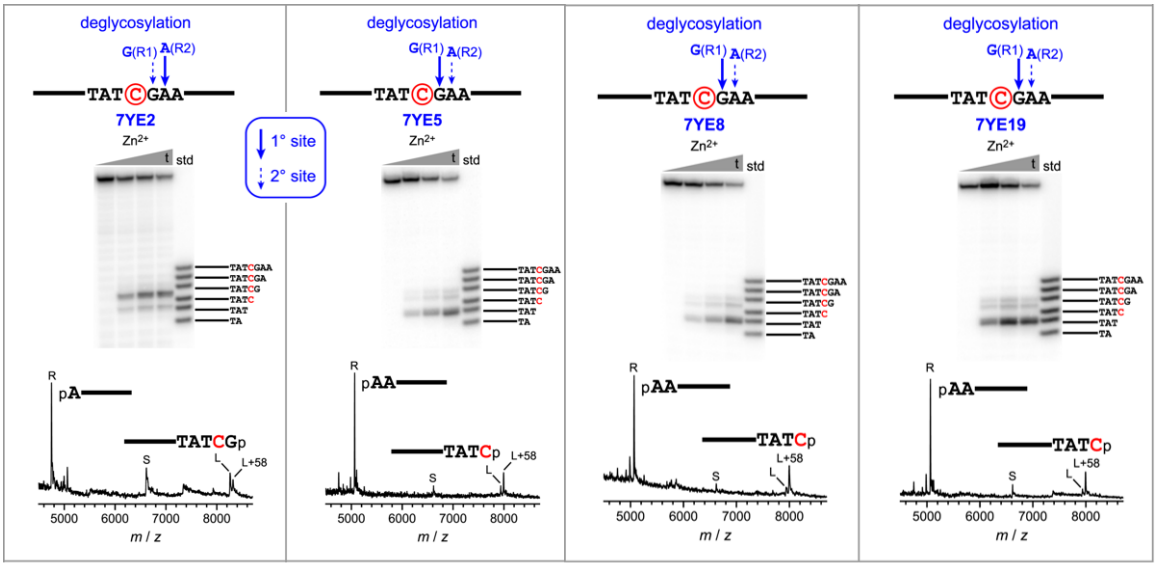

Figure 2.26: Assays of 7YE2, 7YE5, 7YE8, 7YE19 deoxyribozymes that catalyze DNA deglycosylation and β-elimination from the YE selection experiment that used Zn^{2+} (t = 30 s, 30 min, 2 h, 16 h). The arrows above each sequence mark the site(s) of deglycosylation, as assigned on the basis of the MALDI mass spectrometry data. Nucleotides are numbered as part of either the L (left) or R (right) portion of the substrate, counting outward. See Table 2.3 for all mass spectrometry data values. Mass spectrometry peaks labeled S correspond to uncleaved substrate (*z* = 2). In all cases, the assigned deglycosylation sites were consistent with the PAGE standard ladders. Definitive assignment of reaction products was made on the basis of the mass spectrometry data.

All deoxyribozymes identified from the selection with Zn^{2+} and Mg^{2+} have the same deglycosylation site specificities but vary in terms of site preference (Fig. 2.27). These DNA catalysts were evaluated with Zn^{2+} alone, with Zn^{2+} + Mg^{2+} and with Zn^{2+} + Ca^{2+}. 6YF2, 6YF 15 and 6YF 27 are more active with Zn^{2+} + Mg^{2+} than with Zn^{2+} alone or Zn^{2+} + Ca^{2+}. In contrast, 6YF13, 6YF 16 and 6YF 20 have nearly equivalent activity with Zn^{2+} alone and Zn^{2+} + Mg^{2+}, and they are slightly suppressed by Ca^{2+} (Fig. 2.27).

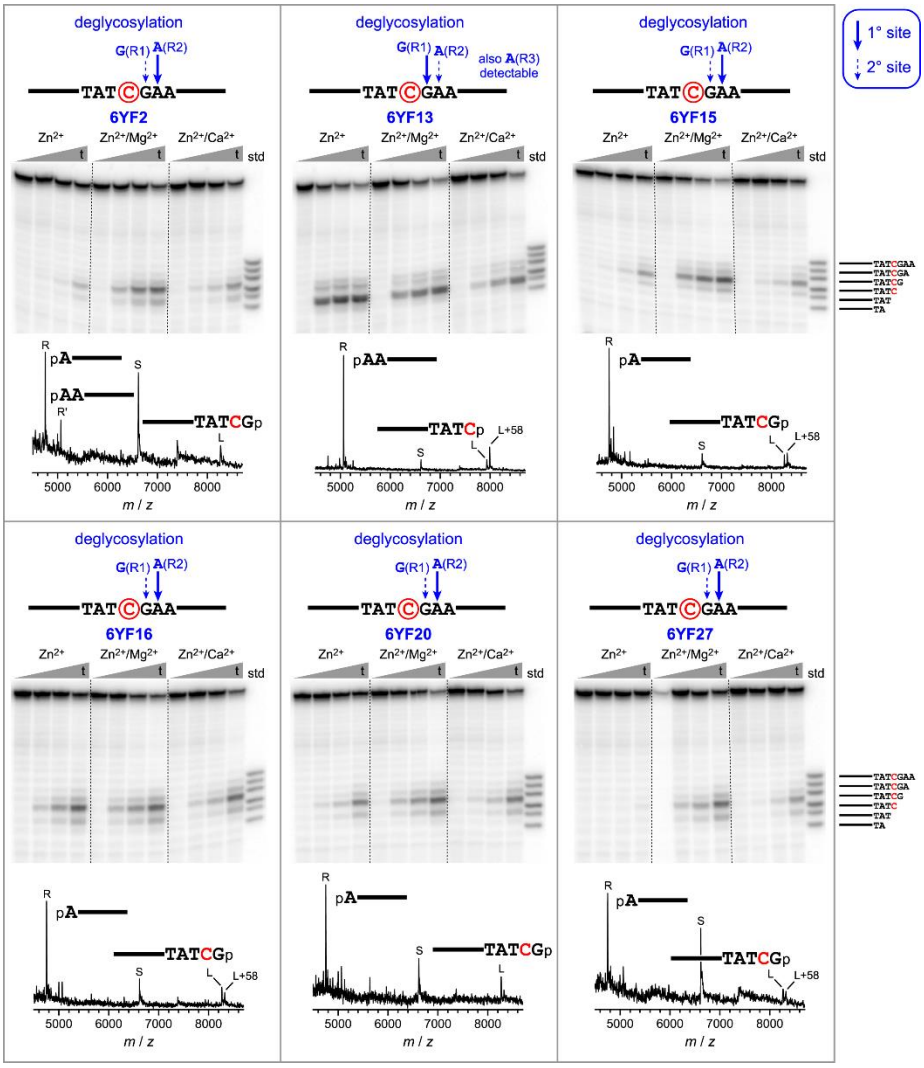

Figure 2.27: Assays of 6YF2, 13, 15, 16, 20, 27 deoxyribozymes that catalyze DNA deglycosylation from the selection experiment that used Zn^{2+} + Mg^{2+} (t = 30 s, 30 min, 2 h, 16 h). The arrows above each sequence mark the site(s) of deglycosylation, as assigned on the basis of the MALDI mass spectrometry data. See Table 2.3 for all mass spectrometry data values (for products formed using Zn^{2+} + Mg^{2+}). In all cases, the assigned deglycosylation sites were consistent with the PAGE standard ladders. Definitive assignment of reaction products was made on the basis of the mass spectrometry data.

Deoxyribozymes identified from the selection with Zn^{2+} and Ca^{2+} display a range of deglycosylation site specificities and site preferences, with one example, 8YG24, equally active towards both the unpaired C and its adjacent T nucleobase. 8YG1 and 8YG7 catalyze depyrimidination of the same nucleobases as 8YG24. 8YG7 depurinates the adjacent A nucleobases to the right of the unpaired nucleotide. These DNA catalysts were tested under selection conditions and with Zn^{2+} alone as well as with Zn^{2+} + Mg^{2+}. 8YG1 and 8YG24 clearly require Zn^{2+} + Ca^{2+} for activity. 8YG7 is more active with Zn^{2+} + Mg^{2+} than with Zn^{2+} alone or Zn^{2+} + Ca^{2+}. 8YG29 is inactive with Zn^{2+} alone and approximately equally active with Zn^{2+} + Mg^{2+} and Zn^{2+} + Ca^{2+} (Fig. 2.28).

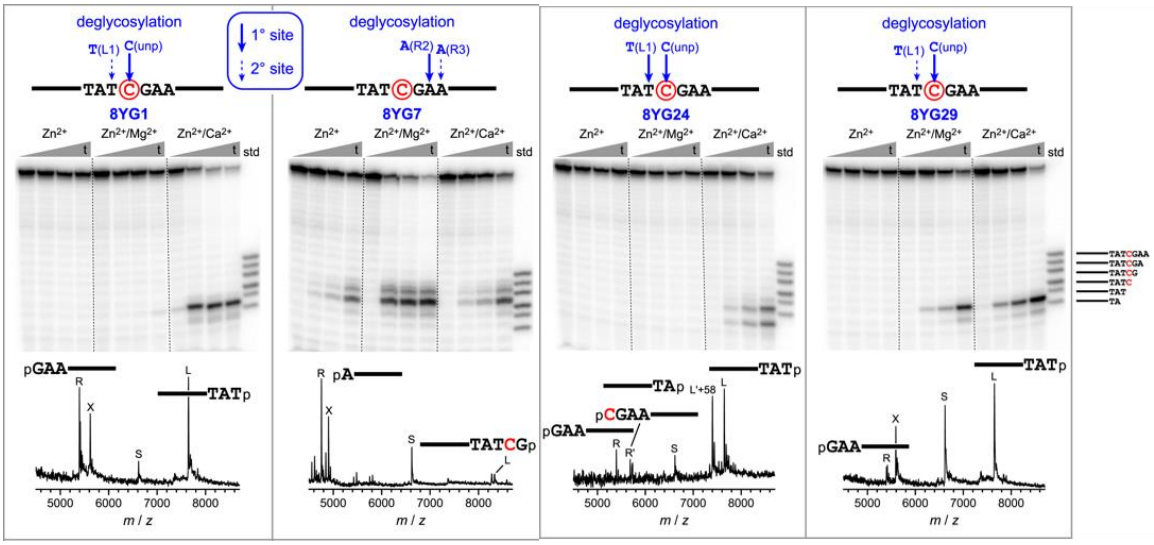

Figure 2.28: Assays of 8YG1, 8YG7, 8YG24 and 8YG29 deoxyribozymes that catalyze DNA deglycosylation and β-elimination from the selection experiment that used Zn^{2+} + Ca^{2+} (t = 30 s, 30 min, 2 h, 16 h). The arrows above each sequence mark the site(s) of deglycosylation, as assigned on the basis of the MALDI mass spectrometry data. Nucleotides are numbered as part of either the L (left) or R (right) portion of the substrate, counting outward. See Table 2.3 for all mass spectrometry data values (for products formed using Zn^{2+} + Ca^{2+}). Mass spectrometry peaks labeled S correspond to uncleaved substrate (*z* = 2); mass spectrometry peaks labeled X were unassigned and not reproducibly observed. In all cases, the assigned deglycosylation sites were consistent with the PAGE standard ladders. Definitive assignment of reaction products was made on the basis of the mass spectrometry data.

deoxy-ribozyme	deglycos. site(s) [a]	mass L calcd.	mass L found	L error, % (found – calcd.)	mass R calcd.	mass R found	R error, % (found – calcd.)
8YM4	G(R1): L, R	7937.2	7936.2	-0.01	5065.3	5065.0	-0.01
8YM17	T(L1): L, R	7343.8	7343.8	0	5683.7	5682.8	-0.02
	A(L2): L′, R′	7030.6	7029.9	-0.01	5987.9	5987.8	-0.002
8YM20	G(R1): L, R	7937.2	7940.9	+0.05	5065.3	5067.2	+0.04
	A(R2): L′, R′	8266.4	not obs.	—	4752.1	4754.2	+0.04
	A(R3): L″, R″	8579.6	8574.5	-0.06	4438.9	4441.1	+0.05
8YM26	T(L1): L, R	7343.8	7347.2	+0.05	5683.7	5686.5	+0.05
	A(L2): L′, R′	7030.6	7034.0	+0.05	5987.9	5991.6	+0.06
7YE2	A(R2): L, R	8266.4	8262.6	-0.05	4752.1	4749.9	-0.05
	G(R1): L′, R′	7937.2	not obs.	—	5065.3	5062.6	-0.05
7YE5	G(R1): L, R [b]	7995.2	7993.9	-0.02	5065.3	5064.8	-0.01
7YE8	G(R1): L, R [b]	7995.2	7997.4	+0.03	5065.3	5064.3	-0.02
7YE19	G(R1): L, R [b]	7995.2	7994.7	-0.01	5065.3	5065.7	+0.01
6YF2	A(R2): L, R	8266.4	8267.6	+0.01	4752.1	4752.8	+0.01
	G(R1): L′, R′	7937.2	not obs.	—	5065.3	5066.5	+0.02
6YF13	G(R1): L, R [b]	7995.2	7992.2	-0.04	5065.3	5064.4	-0.02
6YF15	A(R2): L, R	8266.4	8258.9	-0.09	4752.1	4747.4	-0.10
6YF16	A(R2): L, R	8266.4	8266.2	-0.002	4752.1	4752.7	+0.01
	G(R1): L′, R′ [b]	7995.2	not obs.	—	5065.3	5066.1	+0.02
6YF20	A(R2): L, R	8266.4	8268.5	+0.03	4752.1	4753.2	+0.02
6YF27	A(R2): L, R	8266.4	8264.4	-0.02	4752.1	4750.8	-0.03
8YG1	C(unp): L, R	7648.0	7650.6	+0.03	5394.5	5396.9	+0.04
8YG7	A(R2): L, R [b]	8324.4	8334.2	+0.12	4752.1	4756.9	+0.10
8YG24	C(unp): L, R	7648.0	7648.6	+0.01	5394.5	5394.5	0
	T(L1): L′, R′ [b]	7401.8	7402.0	+0.003	5683.7	5684.1	+0.01
8YG29	C(unp): L, R	7648.0	7650.8	+0.04	5394.5	5396.5	+0.04

Table 2.3: Mass spectrometry assays of cleavage products for DNA-deglycosylating deoxyribozymes.

[a] See Fig. 2.25 (8YM deoxyribozymes, Mn^{2+}), Fig. 2.26 (7YE deoxyribozymes, Zn^{2+}), Fig. 2.27 (6YF deoxyribozymes, Zn^{2+}/Mg^{2+}) and Fig. 2.28 (8YG deoxyribozymes, Zn^{2+}/Ca^{2+}) for assigned deglycosylation site(s). Some of the secondary deglycosylation sites observed by PAGE were sufficiently weak such that no corresponding product peaks could be observed by mass spectrometry; a row is omitted altogether for these secondary sites. In addition, the L peaks for several of the secondary sites were not observed, although the R peaks were observed; these instances are denoted by "not obs." in the table.

[b] In this case, the L product peak was observed primarily in the L+58 form. The L calcd. value was computed from the initially expected L value by adding the mass of $C_2H_2O_2$.

2.3 Summary

In this chapter, in vitro selection, an artificial evolution process, was used to identify single-stranded DNA-cleaving deoxyribozymes starting from a pool of random sequences. We evaluated the ability of numerous polyvalent metal ions at different concentrations (ranging from micromolar to milimolar) to support catalytic activity of DNA enzymes. Due to several unforeseen complications during this work, improvements to the in vitro selection design were made.

Initially, an existing DNA-hydrolyzing deoxyribozyme was tested in the presence of trivalent lanthanide ions. Although it was concluded that these metal ions cannot support catalysis by this deoxyribozyme, an important discovery was made. Zn^{2+} was shown to be the sole required polyvalent metal ion cofactor for this DNA catalyst initially believed to also require Mn^{2+} for its activity. In subsequent work, contamination by catalytic DNA sequences found in earlier efforts subverted the process of new deoxyribozyme identification via in vitro selection. Full redesign of one primer-binding region still allowed for contaminating sequences to propagate. Only substantial modification of both fixed substrate-binding sequences prevented contamination. This principle has been applied in all subsequent work in the Silverman laboratory.

Using the new methodology, new selection experiments were initiated. The key selection step was performed in the presence of 10 μM lanthanide ion (Ce^{3+}, Eu^{3+}, or Yb^{3+}) along with 1 mM Zn^{2+}. We identified numerous DNA-hydrolyzing deoxyribozymes, most of which strictly require the lanthanide ion as well as Zn^{2+} for catalytic activity. These DNA catalysts have a range of lanthanide dependences, including some deoxyribozymes that strongly favor one particular lanthanide ion. Typically, a DNA catalyst's preferred

lanthanide was present in the selection experiment from which this catalyst was identified. DNA enzymes that function well with more than one lanthanide ion were also found. Intriguingly, two of the Yb^{3+}-dependent deoxyribozymes function well with Yb^{3+} alone ($K_{d,app}$ ~10 μM, in the absence of Zn^{2+}) and have little or no activity with Eu^{3+} or Ce^{3+}. In contrast to these selection outcomes when lanthanide ions were present, new selections with Zn^{2+} or Mn^{2+} alone, or Zn^{2+} with Mg^{2+}/Ca^{2+}, led primarily to deoxyribozymes that cleave single-stranded DNA by deglycosylation and β-elimination rather than by hydrolysis, including several instances of depyrimidination.

2.3.1 Implications for future selection efforts

The two orders of magnitude concentration difference between 10 μM lanthanide and 1 mM Zn^{2+} may enable new applications of deoxyribozymes, particularly in contexts where high polyvalent metal ion concentrations must be avoided. Future efforts should pursue direct identification of new deoxyribozymes that require only lanthanide ions by including lanthanides as the sole polyvalent metal ions present during the selection process. Efforts that include increased time pressure for faster reactivity may also be beneficial, considering that most of the new deoxyribozymes have half-lives on the order of several hours, although the fastest lanthanide-dependent deoxyribozymes have half-lives of ca. 0.5 h (Table 2.1). In general, on the basis of our overall findings, lanthanide ions deserve closer attention as metal ion cofactors when developing new nucleic acid catalysts. For the DNA-cleaving deoxyribozymes that were identified by selection in the presence of Mn^{2+} alone, Zn^{2+} alone, or Zn^{2+} with Mg^{2+}/Ca^{2+}, deglycosylation was primarily observed.

Several instances DNA-catalyzed depyrimidination further highlight the catalytic competence of DNA.

2.4 Materials and Methods

2.4.1 Oligonucleotides and in vitro selections

Oligonucleotides were prepared by solid-phase synthesis at Integrated DNA Technologies (Coralville, IA) and purified by denaturing 20% or 8% PAGE. To avoid unwanted chelation of Zn^{2+} or lanthanide ions by adventitious EDTA, all substrates used in kinetic assays were extracted from gels using TN buffer (10 mM Tris, pH 8.0, 300 mM NaCl) lacking EDTA and precipitated with ethanol. The DNA substrate sequences during the selection process were 5′-CTACCTTTATGCGTATCGAAGGAGGCTTTCGgga-3′, where the 3′-terminal gga residues were RNA rather than DNA to enable ligation of the substrate to the deoxyribozyme pool strand by T4 RNA ligase during each selection round. The DNA substrate sequences used for cleavage assays were the same as those used during selection, except the 3′-terminal –Ggga was replaced with –GGGA (i.e., all DNA). The deoxyribozyme pool strand during selections was 5′-CGAA<u>CGAAAGCCTCCTTC</u>-N_{40}-<u>ATACGCATAAAGGTAG</u>AGCTGATCCTGATGG-3′, where the two underlined regions denote the substrate binding arms. The 5′-CGAA was replaced with 5′-CC for individual deoxyribozymes prepared by solid-phase synthesis and used in the single-turnover in trans assays. The two PCR primers used during selection were 5′-CGAACGAAAGCCTCCTTC-3′ (5′-phosphorylated to allow ligation of the product by T4 RNA ligase) and 5′-$(AAC)_4$XCCATCAGGATCAGCT-3′ (where X denotes Glen

Spacer 18, which is a PEG spacer that stops extension by Taq polymerase and leads to a size difference between the two PCR product strands).

The selection experiments were performed using 200 pmol (~10^{14} molecules) of DNA in each initial round of selection, where sequence space encompasses $4^{40} \approx 10^{24}$ possibilities (10^{-10} sampling of sequence space). During the key DNA hydrolysis step of each selection round, the sample was incubated at 37 °C for 12 h. Samples that included Zn^{2+} and a lanthanide ion contained 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 10 μM $LnCl_3$ and 150 mM NaCl, where Ln = Ce, Eu, or Yb, in the form of $CeCl_3 \bullet 7H_2O$ (Aldrich 99.9%), $EuCl_3 \bullet 6H_2O$ (Acros 99.99%), or $YbCl_3 \bullet 6H_2O$ (GFS 99.99%). Lanthanide ion stocks were 100 mM $LnCl_3$ and 200 mM HNO_3, diluted with H_2O to 5 × relative to the final desired concentration (due to the low final lanthanide ion concentration, the small amount of HNO_3 contributed to the final sample was negligible relative to the much higher HEPES buffer concentration). Samples that included Zn^{2+} either alone or with Mg^{2+}/Ca^{2+} contained 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 40 mM $MgCl_2$ or $CaCl_2$ if appropriate and 150 mM NaCl. The Zn^{2+} was added from a 10 × stock of 10 mM $ZnCl_2$ in 20 mM HNO_3 and 200 mM HEPES, pH 7.5, this 10 × stock was itself prepared from a solution of 100 mM $ZnCl_2$ in 200 mM HNO_3 and a solution of 1 M HEPES, pH 7.5.[32] Samples that included Mn^{2+} contained 50 mM HEPES, pH 7.5, 20 mM $MnCl_2$ and 150 mM NaCl.

Cloning and initial screening of individual deoxyribozyme clones was performed using DNA strands prepared by PCR from miniprep DNA derived from individual E. coli colonies. The miniprep DNA samples were first assayed by digestion with EcoRI to ascertain the presence of the expected insert. The concentration of each PAGE-purified deoxyribozyme strand was estimated from the intensity (UV shadowing) relative to

suitable standards. Each screening assay used ~ 0.1 pmol of 5′-^{32}P-radiolabeled L substrate and 2 pmol of the deoxyribozyme strand.

2.4.2 DNA cleavage assay procedures

The DNA cleavage assays using individual deoxyribozymes were performed under single-turnover in trans conditions using a DNA substrate with sequence 5′-CTACCTTTATGCGTATCGAAGGAGGCTTTCGGGA-3′ (i.e. all-DNA sequence version of the substrate used during the selection experiments). The 5′-^{32}P-radiolabeled substrate (S) was the limiting reagent relative to the deoxyribozyme (E). A sample containing 0.2 pmol of S and 20 pmol of E was annealed in 5 mM HEPES, pH 7.5 (or other buffer as appropriate), 15 mM NaCl and 0.1 mM EDTA (EDTA was omitted in all assays with lanthanides) by heating at 95 °C for 3 min, cooling on ice for 5 min and heating at 37 °C for 2 min. The cleavage reaction was initiated using the same method as for selections described in the previous section. Final concentrations for all single-turnover experiments were 10 nM S and 1 μM E in 10-25 μL total volume. The metal ion stocks were added last to the final sample, which was divided into 2 μL aliquots that were all incubated at 37 °C. At appropriate times, aliquots were quenched with 5 μL of stop solution [80% formamide, TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3), 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol]. Samples were separated by 20% PAGE and quantified with a PhosphorImager. The final incubation conditions were as described above for the selection experiments. The lanthanide ion concentration was 3 to 80 μM. When only lanthanide ions were present (i.e., Zn^{2+} was absent), the HEPES

concentration was 50 mM. Assay samples were separated by 20% PAGE and quantified with a PhosphorImager. Values of k_{obs} and final yield were obtained by fitting the data directly to first-order kinetics, i.e.,

$$\text{yield} = Y(1 - e^{-kt}),$$

where $k = k_{obs}$ and Y = final yield. When k_{obs} was sufficiently low such that an exponential fit was not meaningful, the initial points were fit to a straight line and k_{obs} was taken as the slope of the line.

On each assay gel was loaded the same series of oligonucleotide standards. These standards provide a common reference point for comparing gel images and also provided a basis for preliminary assignment of reaction products, albeit with some ambiguity due to near-coincident migration of (i) oligonucleotides of length n with 3′-hydroxyl and (ii) oligonucleotides of length n+1 with 3′-phosphate (greater mass reduces the migration rate, whereas greater charge increases the migration rate). Because each standard oligonucleotide was synthesized with 3′-hydroxyl rather than 3′-phosphate, there is a vertical offset for each 3′-phosphate product relative to the corresponding 3′-hydroxyl standard band (the product migrates faster than the standard). The definitive assignment of each deoxyribozyme's reaction product was made on the basis of the mass spectrometry data.

2.4.3 Mass spectrometry

Samples for mass spectrometry were prepared as follows. A 100 μL sample containing 60 pmol of substrate and 200 pmol of deoxyribozyme was annealed in 5 mM

HEPES, pH 7.5, 15 mM NaCl and 0.1 mM EDTA (the EDTA was omitted for lanthanide-containing reactions) by heating at 95 °C for 3 min and cooling on ice for 5 min. The cleavage reaction was initiated by addition of stock solutions to a final volume of 200 µL containing HEPES, pH 7.5 and metal ions as described above for in vitro selection. The sample was incubated at 37 °C for 16 h. The nucleic acids were precipitated by adding 20 µL of 3 M NaCl and 660 µL of ethanol and dissolved in 40 µL water. The sample was desalted in 10 µL portions using four C_{18} ZipTips and analyzed by MALDI mass spectrometry with 3-hydroxypicolinic acid as matrix. All MALDI mass spectra were obtained in the mass spectrometry laboratory of the UIUC School of Chemical Sciences.

2.5 References

1. Silverman, S. K., Deoxyribozymes: selection design and serendipity in the development of DNA catalysts. *Acc. Chem. Res.* **2009,** *42,* 1521-1531.
2. Chandra, M.; Sachdeva, A.; Silverman, S. K., DNA-catalyzed sequence-specific hydrolysis of DNA. *Nat. Chem. Biol.* **2009,** *5*, 718-720.
3. Xiao, Y.; Chandra, M.; Silverman, S. K., Functional compromises among pH tolerance, site specificity, and sequence tolerance for a DNA-hydrolyzing deoxyribozyme. *Biochemistry* **2010,** *49*, 9630-9637.
4. Xiao, Y.; Allen, E. C.; Silverman, S. K., Merely two mutations switch a DNA-hydrolyzing deoxyribozyme from heterobimetallic (Zn^{2+}/Mn^{2+}) to monometallic (Zn^{2+}-only) behavior. *Chem. Commun.* **2011,** *47*, 1749-1751.

5. Velez, T. E.; Singh, J.; Xiao, Y.; Allen, E. C.; Wong, O. Y.; Chandra, M.; Kwon, S. C.; Silverman, S. K., Systematic evaluation of the dependence of deoxyribozyme catalysis on random region length. *ACS Comb. Sci.* **2012,** *14*, 680-687.
6. Burrows, C. J.; Muller, J. G., Oxidative Nucleobase Modifications Leading to Strand Scission. *Chem. Rev.* **1998,** *98*, 1109-1152.
7. Balasubramanian, B.; Pogozelski, W. K.; Tullius, T. D., DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone. *Proc. Natl. Acad. Sci. U.S.A.* **1998,** *95*, 9738-9743.
8. Greenberg, M. M., Elucidating DNA damage and repair processes by independently generating reactive and metastable intermediates. *Org. Biomol. Chem.* **2007,** *5*, 18-30.
9. Gates, K. S., An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem. Res. Toxicol.* **2009,** *22*, 1747-1760.
10. Bunzli, J. C.; Piguet, C., Taking advantage of luminescent lanthanide ions. *Chem. Soc. Rev.* **2005,** *34*, 1048-1077.
11. Geyer, C. R.; Sen, D., Lanthanide probes for a phosphodiester-cleaving, lead-dependent, DNAzyme. *J. Mol. Biol.* **1998,** *275*, 483-489.
12. Feig, A. L.; Panek, M.; Horrocks, W. D., Jr.; Uhlenbeck, O. C., Probing the binding of Tb(III) and Eu(III) to the hammerhead ribozyme using luminescence spectroscopy. *Chem. Biol.* **1999,** *6*, 801-810.
13. Komiyama, M.; Takeda, N.; Shigekawa, H., Hydrolysis of DNA and RNA by lanthanide ions: mechanistic studies leading to new applications. *Chem. Commun.* **1999**, *35*, 1443-1451.

14. Walter, N. G.; Yang, N.; Burke, J. M., Probing non-selective cation binding in the hairpin ribozyme with Tb(III). *J. Mol. Biol.* **2000,** *298*, 539-555.

15. Mundoma, C.; Greenbaum, N. L., Sequestering of Eu(III) by a GAAA RNA tetraloop. *J. Am. Chem. Soc.* **2002,** *124*, 3525-3532.

16. Sheppard, T. L.; Ordoukhanian, P.; Joyce, G. F., A DNA enzyme with N-glycosylase activity. *Proc. Natl. Acad. Sci. U.S.A.* **2000,** *97*, 7802-7807.

17. Nowakowski, J.; Shim, P. J.; Prasad, G. S.; Stout, C. D.; Joyce, G. F., Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme. *Nat. Struct. Biol.* **1999,** *6*, 151-156.

18. Breaker, R. R.; Joyce, G. F., A DNA enzyme that cleaves RNA. *Chem. Biol.* **1994,** *1*, 223-229.

19. Greenwood, N. N.; Earnshaw, A., *Chemistry of the Elements* (2nd Ed.). Butterworth-Heinemann: Oxform, **1997**; 1233.

20. Zoltewicz, J. A.; Clark, D. F.; Sharpless, T. W.; Grahe, G., Kinetics and mechanism of the acid-catalyzed hydrolysis of some purine nucleosides. *J. Am. Chem. Soc.* **1970,** *92*, 1741-1749.

21. Berti, P. J.; McCann, J. A., Toward a detailed understanding of base excision repair enzymes: transition state and mechanistic analyses of N-glycoside hydrolysis and N-glycoside transfer. *Chem. Rev.* **2006,** *106*, 506-555.

22. Schroeder, G. K.; Wolfenden, R., Rates of spontaneous disintegration of DNA and the rate enhancements produced by DNA glycosylases and deaminases. *Biochemistry* **2007,** *46*, 13638-13647.

23. Muller, S. J.; Caradonna, S., Isolation and characterization of a human cDNA encoding uracil-DNA glycosylase. *Biochim. Biophys. Acta.* **1991,** *1088*, 197-207.

24. Olsen, L. C.; Aasland, R.; Wittwer, C. U.; Krokan, H. E.; Helland, D. E., Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. *EMBO J* **1989,** *8*, 3121-3125.

25. Lindahl, T.; Ljungquist, S.; Siegert, W.; Nyberg, B.; Sperens, B., DNA N-glycosidases: properties of uracil-DNA glycosidase from Escherichia coli. *J. Biol. Chem.* **1977,** *252*, 3286-3294.

26. Mol, C. D.; Arvai, A. S.; Slupphaug, G.; Kavli, B.; Alseth, I.; Krokan, H. E.; Tainer, J. A., Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell* **1995,** *80*, 869-878.

27. Friedman, J. I.; Stivers, J. T., Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry* **2010,** *49*, 4957-4967.

28. Dalhus, B.; Laerdahl, J. K.; Backe, P. H.; Bjoras, M., DNA base repair – recognition and initiation of catalysis. *FEMS Microbiol. Rev.* **2009,** *33*, 1044-1078.

29. Mok, Y. G.; Uzawa, R.; Lee, J.; Weiner, G. M.; Eichman, B. F.; Fischer, R. L.; Huh, J. H., Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107*, 19225-19230.

30. Pogozelski, W. K.; Tullius, T. D., Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety. *Chem. Rev.* **1998,** *98*, 1089-1108.

31. Joshi, R. R.; Likhite, S. M.; Kumar, R. K.; Ganesh, K. N., DNA cleavage by Cu(II)-desferal: identification of C1'-hydroxylation as the initial event for DNA damage. *Biochim. Biophys. Acta.* **1994,** *1199,* 285-292.

32. Hoadley, K. A.; Purtha, W. E.; Wolf, A. C.; Flynn-Charlebois, A.; Silverman, S. K., Zn^{2+}-Dependent Deoxyribozymes That Form Natural and Unnatural RNA Linkages. *Biochemistry* **2005,** *44,* 9217-9231.

Chapter 3: Efforts to Identify Novel DNA Aptamers of ATP

3.1 Introduction

The term "aptamer" is derived from the Latin word *aptus* (to fit) and from the Greek word *meros* (part).[1] Single-stranded DNA and RNA sequences that bind small molecules are called aptamers. These functional nucleic acids existed in nature since the early stages of evolution of life on Earth. For example, a self-cleaving ribozyme described in Chapter 1 of this thesis specifically binds a free molecule of GTP, a necessary cofactor for its intron-splicing mechanism. Riboswitches, also discussed in this thesis, possess an integral aptamer region, which upon binding its ligand, either inhibits or activates gene expression inside a cell. According to the RNA World hypothesis, nucleic acid aptamers (in addition to other functional nucleic acids) played a far greater role a billion years ago than in today's protein-based life. And although the binding (also, catalytic) functions in nature have mostly been replaced by polypeptides, the underlying principle of specific sequence-dependent nucleic acid interactions with molecular targets still holds true.

In 1990, two laboratories independently identified artificial RNA aptamers.[1,2] The Szostak lab from the Massachusetts General Hospital generated this kind of binders for numerous organic dyes and coined the term "in vitro selection"; the Gold lab from the University of Colorado identified RNA aptamers for T4 DNA polymerase, giving the name SELEX for the process of selection. The letter "L" in this acronym stands for "ligand" and this term should be applied specifically to artificial evolution of nucleic acid binders, not catalysts.[3] The SELEX process is similar to the previously described method of artificial evolution to obtain novel ribozymes (for RNA) and deoxyribozymes (for DNA). It is a

matter of personal preference which nucleic acid to use in SELEX; however, both DNA versions of RNA aptamers and RNA versions of DNA aptamers are inactive.[4] Synthesis of the random sequence pool and its general design with the variable segment flanked by constant primer-binding regions and amplification by PCR is identical for both processes. Instead of incubation with substrates, during the key selection step of the SELEX process nucleic acid sequences are exposed to molecular targets that are covalently attached to a solid-phase support (e.g., agarose beads). Incubation conditions for this step must be chosen carefully, often with regard to likely future use for possible aptamers in mind. Variables include buffer identity and pH, temperature and the time of incubation as well as metal ions and their concentrations. Sequences with high affinity towards this molecular target bind the small molecule and become immobilized while the rest of the sequences are eluted out and discarded. Single-stranded nucleic acid sequences from the pool bound to the immobilized targets undergo another incubation under the same conditions but in the presence of free target molecules. For more stringent selections, the concentration of free target molecules should be decreased in later rounds. Sequences that switch from binding immobilized targets to binding free targets can then be eluted and amplified. In other cases, free targets are not used. Instead, the mixture is incubated with denaturing or chelating agents to break interactions between DNA and immobilized target molecules. After the sequence pool becomes enriched with binders of desired affinity, individual sequences are cloned and studied to determine parameters such as binding constants. Aptamers for large molecules (specifically, proteins) can be identified via the routes described above. In addition to these methods, proteins with net positive charge (which is true of most proteins) can be separated on filters made of nitrocellulose, which are negatively charged.

Aptamers have found numerous uses in today's biomedical and biotechnological industries. They are applied in food safety applications, as biosensors for molecular diagnostics, as drug delivery agents or as probes for flow cytometry.[5-8] In this work we are interested in the modular incorporation of a DNA aptamer sequence into a deoxyribozyme. This work is described in Chapter 4 of this thesis.

Examples of some of the best-characterized aptamers include RNA and DNA aptamers that were selected to bind ATP (adenosine triphosphate) molecules, first published by the Szostak lab in 1993 and 1995, respectively.[9,10] Due to our laboratory's interest in functional DNA (rather than RNA) molecules, the DNA aptamer of ATP was of interest for this work. It was identified from a N_{72} random sequence pool and in follow-up experiments, a 27-nucleotide motif was shown to be responsible for binding two ATP molecules.[10] Via the method of equilibrium filtration, this aptamer was found to bind adenosine with an estimated K_d of about 6 ± 3 μM. Extensive mutagenesis studies were also performed on the aptamer sequence to identify key nucleotides involved in the binding of free ATP molecules. Additionally, numerous analogs of ATP were used to determine specificity of the target–aptamer interaction. These experiments led to an accurate view of the structure that these sequences adapt when ATP molecules are bound. The results were later confirmed by NMR studies of the complexes in solution.[11]

Encouraged by the ability to discover aptamers via SELEX and, specifically, DNA aptamers of ATP by other scientists,[10,12] we chose to identify and characterize new ATP aptamers followed by their modular incorporation into the design of deoxyribozyme selections. We also planned to use the Szostak ATP aptamer in these future selection experiments either alongside the newly identified aptamers or as the sole aptamer motif in

the case of failing to find ATP aptamers in this effort. Because ATP molecules serve as substrates for countless biochemical reactions, we envisioned incorporating an invariable ATP aptamer sequence near the randomized region of the DNA pool for in vitro selections of DNA catalysts. In all deoxyribozymes, the formerly random regions are required both to catalyze a reaction and to bind substrates. An incorporated aptamer module specifically responsible for binding a small molecule would position it in the vicinity of the random region, providing it with an easy access to the substrate. This approach could lead to the discovery of faster, more numerous deoxyribozymes as the random region would only be tasked with catalysis, but not with substrate binding.

3.2 Results and Discussion

3.2.1 Design and implementation of ATP DNA aptamer selections

ATP molecules immobilized on agarose beads through the γ-phosphate were used in this work (Fig. 3.1). The Szostak ATP aptamer was identified with ATP linked through the C8 atom.

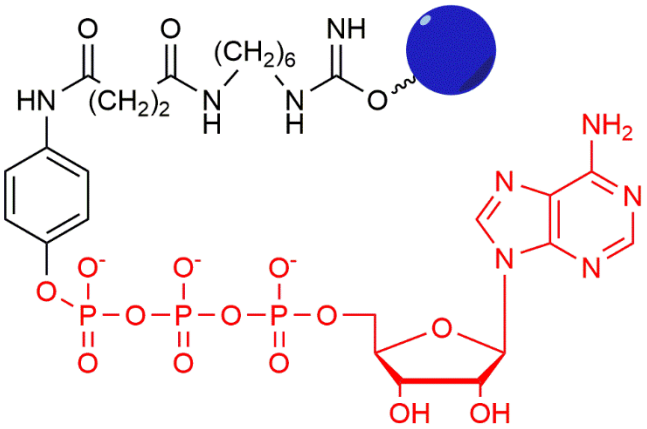

Figure 3.1: ATP molecule immobilized on an agarose bead used for aptamer selections.

Incubations of radiolabeled random pool DNA sequences were performed inside plastic tubes while being nutated to achieve thorough mixing with agarose beads. To elute DNA, the entire mixture was transferred into a column filter (permeable to large DNA molecules but not to much larger agarose beads) and centrifuged. Sequences that did not associate with the beads were collected in the flowthrough. Due to radiolabeling, the amount of single-stranded DNA collected in each round was measured with a scintillation counter.

Single-stranded random pool DNA sequences were incubated with underivatized agarose beads prior to exposing these molecules to the beads with immobilized ATP. This step was very important because some DNA sequences can associate with the solid support itself; these DNA molecules were removed in this step. If this was not done, then such sequences could have outnumbered target-binding DNA molecules and subverted the selection process.

A total of eight selection experiments were initiated, exploring three binary variables (Table 3.1). First, two random pools were designed in a fashion similar to that described in Chapter 2, with the variable region consisting of 30 nucleotides (N_{30}). One of the two pools also included four-nucleotide complementary sequences immediately flanking the random region (B, D, F, H selections). The two primer-binding portions were on the either side of this region. The second pool did not include the additional pairs of nucleotides and its total length was shorter by eight nucleotides (A, C, E, G selections). Random pool designs are presented in Fig. 3.2. How the ATP-bound sequences were eluted constituted the second variable: either free ATP or EDTA (ethylenediaminetetraacetic acid) was used to aid in removal of DNA from solid support. A, B, E and F selections were

eluted with ATP, while C, D, G and H selections were eluted with EDTA. When a sequence dissociates from its target on an agarose bead and binds a free ATP molecule in solution, it can be eluted. EDTA chelates metal ions which are necessary for aptamer structures; removing metal ions would also lead to single-stranded DNA dissociation from agarose-bound ATP. Finally, as the third variable, incubations were performed under two conditions: 40 mM Mg^{2+} with 50 mM CHES, pH 9.0 or 40 mM Mg^{2+}, 20 mM Mn^{2+} with 50 mM HEPES, pH 7.5; all incubations were done with 150 mM NaCl at room temperature.

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Pool	N_{30}	N_{30} + 4 b.p.	N_{30}	N_{30} + 4 b.p.	N_{30}	N_{30} + 4 b.p.	N_{30}	N_{30} + 4 b.p.
Elution	ATP		EDTA		ATP		EDTA	
Metals	Mg^{2+} + Mn^{2+}				Mg^{2+}			

Table 3.1: Letter designation of each aptamer selection experiment and its corresponding random pool design, whether ATP-binding sequences were eluted with ATP or EDTA and the metals used during incubation (50 mM HEPES, pH 7.5 for Mg^{2+} + Mn^{2+} and 50 mM CHES, pH 9.0 for Mn^{2+} both with 150 mM NaCl).

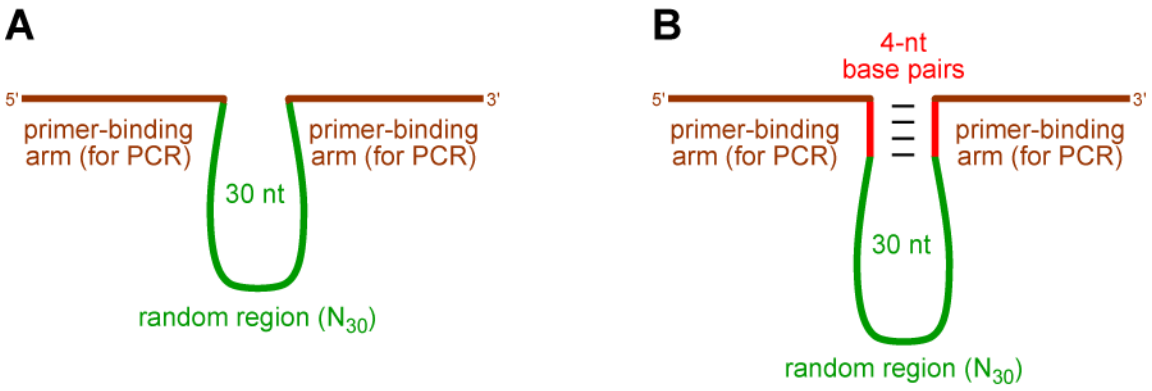

Figure 3.2: Design of the random sequence DNA pools for the ATP aptamer selection experiments. The N_{30} random region was either (A) directly connected to the primer-binding arms in the pool for the A, C, E and G selections or (B) connected via four base-paired nucleotides in the pool for the B, D, F and H selections.

Each of the eight aptamer selections was iterated for ten rounds. Percentage of single-stranded DNA sequences (eluted with either free ATP or with EDTA) relative to the

entire amount of pool molecules eluted is reported in Fig. 3.3. Unlike in the previous selections of DNA catalysts, high backgrounds are observed in activity values for aptamer selections. Only the values above 10-15%, which also increased in the following round, were considered to be real representations of the pool's enrichment with single-stranded ATP binders. The activity of the C selection increased above the background range to over 20% by rounds 7 and 8 (Fig. 3.3). At round 9, all four selections that were incubated in presence of both Mg^{2+} and Mn^{2+} had activities at or above 30%, and individual sequences from these pools were cloned. Even after ten rounds, none of the four selections incubated with Mg^{2+} as the sole polyvalent metal ion had consistent activities above the 10-15% background level, and their progress was terminated.

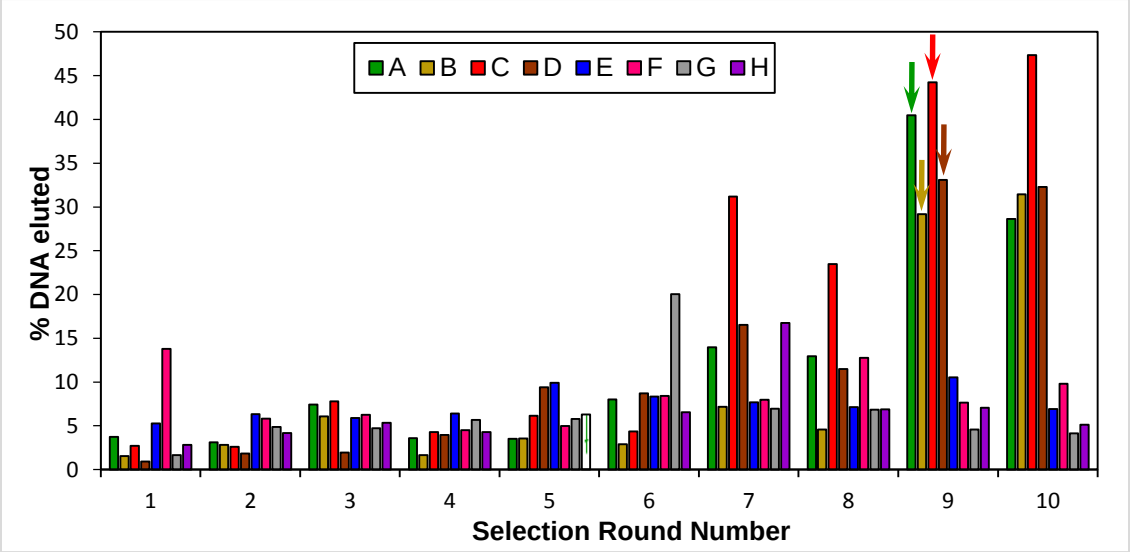

Figure 3.3: Progression of the A-H ATP aptamer selection experiments. For each selection, the round at which individual DNA sequences were cloned is marked with an arrow on the plot.

3.2.2 Characterization of the aptamer candidate sequences

The A, B, C and D ATP aptamer selections were cloned at round 9 and eight *E. coli* colonies were chosen for each selection. Each set of the colonies contained three unique sequences for each of the four pools (Fig. 3.4). All sequences from the B and D selections retained the additional nucleotides flanking the random region although these were not parts of the primer-binding sites and were subject to potential mutations during PCR amplification. Formerly random region lengths also did not change and remained 30 nucleotides long. No regions of conservation among the twelve sequences were observed in the sequence alignment.

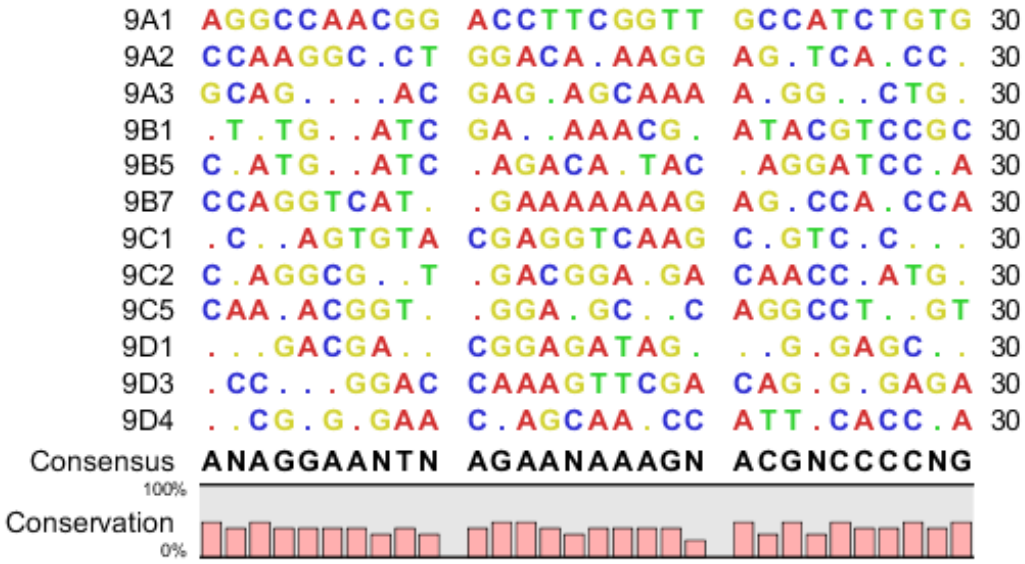

Figure 3.4: Sequences of the initially random (N_{30}) regions of the A-D ATP aptamer selections. All residues identical to those of 9A1 in the corresponding position are depicted as dots.

All twelve aptamer candidate sequences were prepared via solid-phase synthesis for quantitative assessment of their interaction with ATP. All oligonucleotides were synthesized without 5′-binding arms since these regions were not designed to be present

after modular incorporation of aptamer sequences into random pools for DNA catalyst selections. Three methods of ATP binding characterization were employed. First to be tried were bead binding assays in which each sequence was eluted with ATP at increasing concentrations. Next, DMS (dimethyl sulfate) probing experiments were performed. This method is a well-known way to locate the regions of DNA that interact with proteins and other molecules and may also be used to estimate binding strengths.[13] Finally, these aptamer candidate sequences were assayed by equilibrium filtration in which DNA is incubated with free target molecules.[14] Then the mixture is separated on centrifugal filter columns permeable to small ATP molecules but not to large single-stranded DNA. In all three methods, the twelve new sequences were assayed along with two well-characterized ATP aptamers. The first of the two previously known aptamers was already described in section 3.1; it was identified by **D**avid **H**uizenga and Jack Szostak and is referred to as the DH aptamer.[10] The second aptamer is a modification of the first one that binds one ATP molecule instead of two. It is 37 nucleotides in length (10 longer than the DH aptamer) and was identified by **M**ihaela **B**arbu and Milan Stojanovic in 2012. This aptamer is referred to as the MB aptamer.[12]

In bead binding assays, each oligonucleotide was incubated with ATP-derivatized agarose beads and then eluted with free ATP molecules. The elution was done in several steps, each with an increasing concentration of ATP: 30 nM, 100 nM, 300 nM, 1 μM, 3 μM, 10 μM, 30 μM, and finally 100 μM. The total amount of eluted DNA during each

successive step was added to the previous amount and plotted to fit the binding curves using the following equation:

$$Y = 100 \cdot \frac{C^n}{C^n + K_d^n}$$

where Y = % of DNA eluted, C = concentration of ATP and n is the Hill coefficient for ATP binding. Previous studies established that two molecules of ATP bind to the DH aptamer and one molecule of ATP binds to the MB aptamer; therefore, n = 2 and n = 1 values were used for their respective fits. We hypothesized that only one ATP molecule binds to each of the twelve aptamer candidate sequences, so n = 1 was used for these fits. The plots for the two previously identified ATP aptamers (used as controls for this assay) as well as the three 9D aptamer candidate sequences (representative of all twelve aptamer candidates from the A-D aptamer selections) are shown in Fig. 3.5.

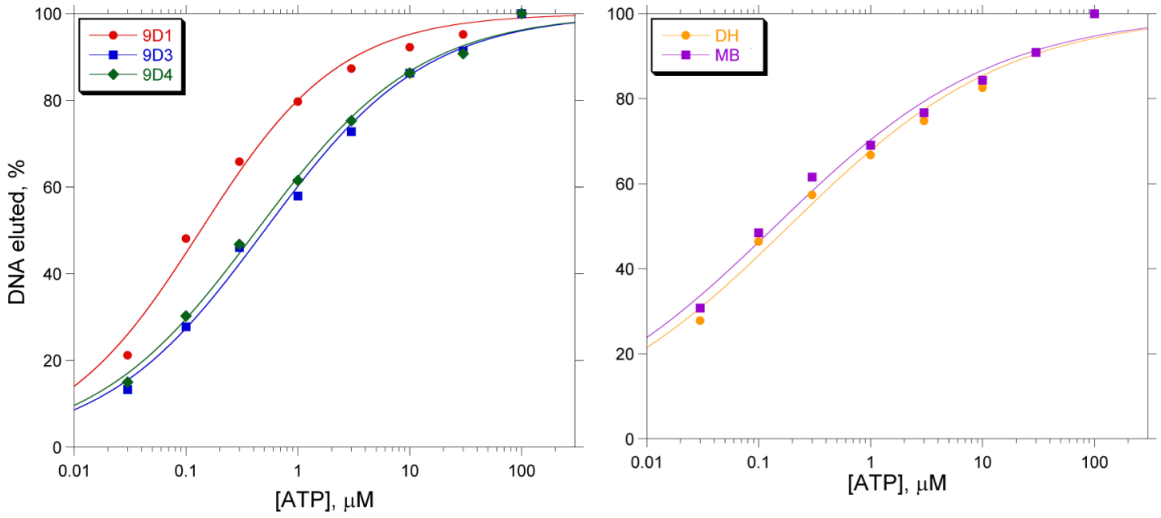

Figure 3.5: Bead-binding assays of 9D1, 9D3 and 9D4 ATP aptamer candidates as well as DH and MB ATP aptamers. Relative amount of radiolabeled sequences eluted with each successive ATP concentration (30 nM to 100 μM) as percent of the total DNA eluted with ATP is plotted.

While the plots fit to an S-shaped curve, reliable aptamer binding data could not be derived from the fits because only as little as 10% of the total amount of sequences eluted with free ATP. The rest of the sequences either remained bound to the beads or were removed when the filters were washed with buffer solutions prior to addition of free ATP, as detected by scintillation counting. Furthermore, bead binding assays do not represent a simple system as both free and immobilized target molecules as well as free and bound (either to an immobilized or a free target) aptamers are present in the mixture. Therefore, we chose to discontinue testing the aptamers by bead binding.

Next, to characterize ATP binding of the twelve aptamer candidates, we turned to DMS probing. These assays are usually used for nucleic acid secondary structure determination, however sometimes they are employed to characterize binding affinities of aptamers and their targets. DMS methylates accessible guanine nucleobases at the N7 position. If a particular guanine of an aptmer candidate is actively interacting with ATP such that its N7 position is blocked, it is not methylated by DMS. The methylated nucleobases are then deglycosylated specifically by piperidine. This glycosidic bond cleavage eventually leads to single-stranded DNA strand scission after two spontaneous β-elimination reactions. After 5′-radiolabeled DNA is treated with DMS followed by piperidine, it is resolved on PAGE as cleavage product strands of different lengths migrate at different rates. Because each cleavage event is the result of a G nucleobase methylation, each accessible guanine can be accounted for by locating the corresponding band on the gel. Any G nucleobases directly involved in ATP binding will not be methylated and a band corresponding to strand cleavage next to these nucleobases will not appear on the gel. At low ATP concentrations, most binding sites in an aptamer are not occupied and all

guanines are accessible for methylation, resulting in the observation of all possible corresponding bands on the gel. At ATP concentrations several-fold higher than the K_d value of an aptamer, such G nucleobases are not methylated and the corresponding bands disappear from the gel.

No previously published DMS probing data for either of the two control aptamers was available. At sufficiently high ATP concentrations, we expected to observe a pattern on PAGE corresponding to no N7 methylation of guanines involved in ATP binding. However, every guanine was observed to be methylated by DMS even at high ATP concentrations, as evidenced by the presence of the corresponding bands on PAGE at all concentrations of ATP tested (Fig. 3.6). Nevertheless, the binding propensity of these two aptamers has been well established, which means that DMS probing simply cannot reveal their interactions with the ATP targets. None of the twelve aptamer candidate sequences exhibited the pattern suggestive of guanine methylation (Fig. 3.6). This finding does not clearly repudiate that all 14 sequences tested are incapable of binding ATP at concentrations tested, as it may be entirely possible for all guanines in an aptamer bound to ATP to be accessible for methylation by DMS.

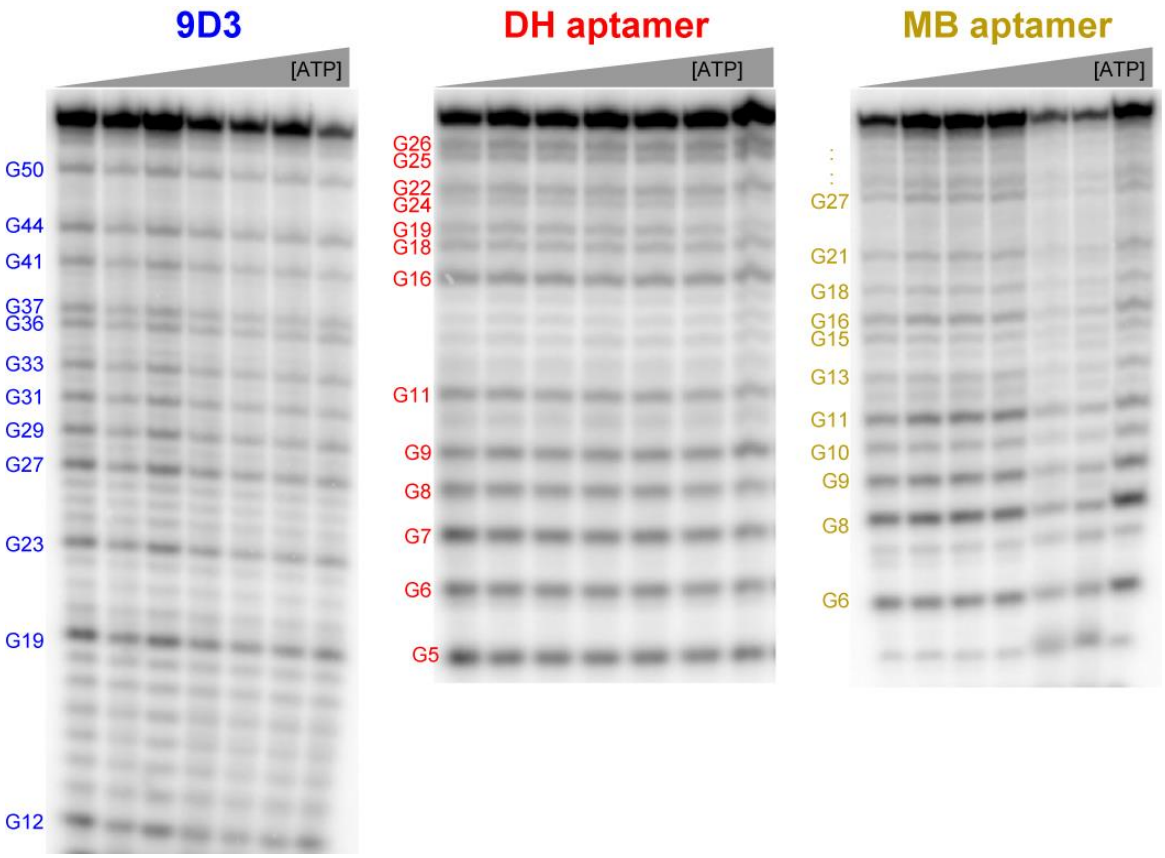

Figure 3.6: PAGE of DMS probing assays of 9D3 ATP aptamer candidate (similar data was observed for the rest of the aptamer candidates, not shown) as well as DH and MB ATP aptamers. ATP concentrations were 0, 1 nM, 10 nM, 100 nM, 1 μM, 10 μM and 100 μM. The more distinct bands represent guanines accessible for methylation by DMS. Positions of G nucleotides in these sequences (presented in the Materials and Methods section) can be confirmed on the gels in a fashion similar to the Sanger sequencing method.

Because we were unable to characterize ATP binding of the twelve aptamer candidates in addition to the two known aptamers by both bead binding assays and DMS probing, we turned to the method of equilibrium filtration. In this technique, filters permeable to small molecules but not to long DNA strands are used. DNA aptamer sequences prevent bound ATP molecules from moving across the filter membrane when the filters are centrifuged. In presence of a DNA aptamer at a sufficient level, ATP concentration in the filtrate should be lower than in the solution prior to centrifugation.

To characterize target-aptamer interactions of the twelve aptamer candidates and the two known ATP aptamers, the ATP concentration was kept constant at 30 μM while the concentration of DNA was varied in the range of 100 nM-100 μM. Trace amounts of radioactive γ-^{32}P-ATP were also added to each sample to measure target concentrations across the membrane (in both filtrate and retentate). Presence of DNA aptamers at high concentrations was expected to lead to appreciable retention of ATP molecules in the retentate solution resulting in a lower number of counts in the filtrate solution. A sample containing a 27-nucleotide sequence incapable of binding ATP was used as reference to compare ATP concentrations in the filtrate. All twelve aptamer candidates and both known ATP aptamers were characterized under (parental) conditions used during selections that led to their identification. Buffer identity, pH and metal ion concentrations used to identify the known ATP aptamers were slightly different than those used in the A-D selections). Even at 100 μM DNA concentration, significant retention of ATP was not observed for any of the twelve aptamer candidates. However, when each known ATP aptamer was present at 10-100 μM, ATP concentration in the filtrate was significantly reduced, indicating successful binding of ATP by the two aptamers, albeit with values of K_d estimated to be several-fold higher than previously reported (Fig. 3.7). For true binding constant measurements, equilibrium filtration assays with even higher but cost-prohibitive aptamer concentrations were not performed. The discrepancy between our estimated K_d values and those previously reported may be due to the method of characterization. In addition to being evaluated under parental (20 mM Tris, pH 7.6, 5 mM Mg^{2+}, 300 mM NaCl) conditions, target binding of the two known ATP aptamers was also characterized under the conditions (50 mM HEPES, pH 7.5, 40 mM Mg^{2+} + 20 mM Mn^{2+}, 150 mM NaCl

and 50 mM CHES, pH 9.0, 40 mM Mg^{2+}, 150 mM NaCl) planned to be used during deoxyribozyme selections with DNA aptamer modules (Figure 3.7). Under the new conditions, both ATP aptamers exhibited levels of ATP binding similar to those under parental conditions. Therefore, due to our inability to identify new ATP binders, we decided to integrate only the previously reported ATP aptamers into the pools for modular selections of DNA enzymes. This work is described in Chapter 4.

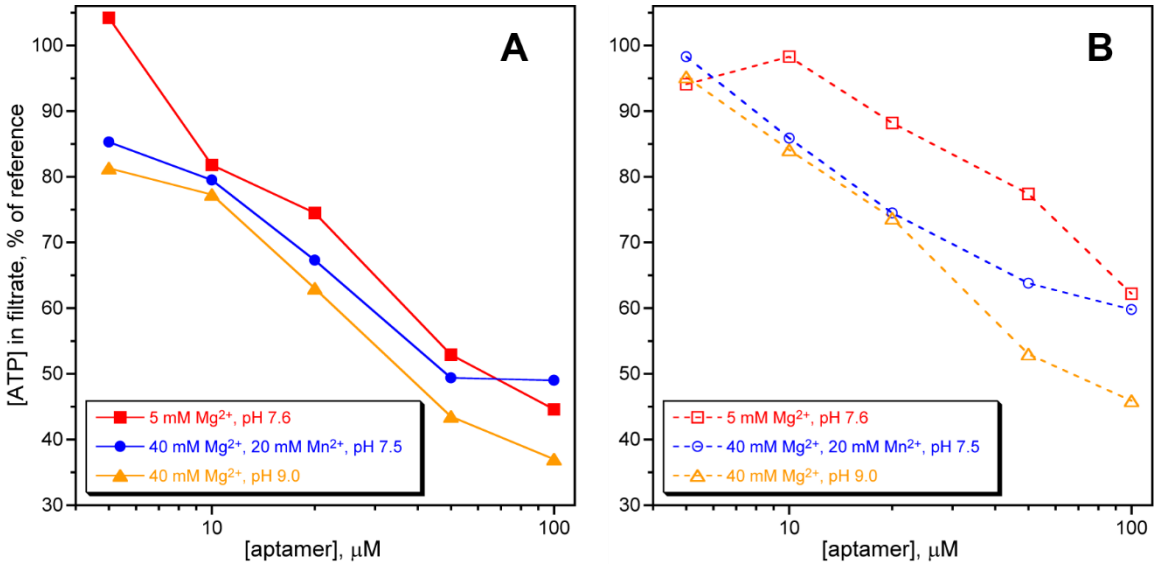

Figure 3.7: Equilibrium filtration assays with 30 μM ATP and 5–100 μM DH (A) or MB (B) aptamer with three different salt/pH conditions. Amount of radioactivity in 20 μL of each filtrate and retentate was measured by scintillation counting and divided by the number of counts in the respective fractions of the reference sample to get the value for the % of reference. This value was directly proportional to ATP concentration due to the presence of radiolabeled ATP molecules in trace amount in each sample.

3.3 Summary

In this chapter, we attempted to identify single-stranded DNA sequences that specifically bind ATP starting from two pools of random sequences via the in vitro selection process. ATP-derivatized agarose beads were used to separate sequences capable of binding targets from the rest of the pool population. Then these sequences were eluted

with either free ATP molecules or with metal-chelating EDTA that interferes with nucleic acid secondary structures, causing dissociation of aptamers from immobilized targets. Selection experiments that were performed with both Mg^{2+} and Mn^{2+} (but not those with Mg^{2+} alone) resulted in domination of pools by sequences that presumably bound to immobilized ATP.

After subsequent cloning and sequencing of twelve individual members of the active pool populations, several attempts to characterize their ATP interactions were made. In addition to aptamer candidates that emerged from these selection experiments, two previously published DNA aptamers for ATP were characterized by the same methods. Assays in which DNA sequences were exposed to agarose-bound ATP molecules and then eluted with increasing concentrations of free ATP did not provide strong evidence of target concentration-specific elution of DNA. This was especially apparent when the two previously published and well-characterized ATP aptamers exhibited the same non-specific elution behavior. Next, aptamer-target complexes were probed with DMS, a molecule that preferably methylates guanine nucleobases at the N7 position when they are not sterically hindered by interactions with bound ATP Unfortunately, all guanine nucleobases were methylated in both previously published ATP aptamers as well as all twelve aptamer candidates. This outcome, however, could not repudiate the binding nature of a DNA sequence and its target, but simply indicated the absence of guanine methylation sites involved in interactions with ATP, at least for the two previously published aptamers. Finally, ATP binding by the two previously published DNA aptamers was confirmed via the method of equilibrium filtration. The same approach did not reveal any evidence of ATP binding by any of the twelve aptamer candidates that emerged from the selection

experiments. Therefore, all twelve aptamer candidates were set aside in favor of the two well-characterized DNA aptamers of ATP to be used in future selection experiments with aptamer modules.

3.4 Materials and Methods

3.4.1 Oligonucleotides and in vitro selection of ATP-binding DNA aptamers

The two random DNA pool sequences used in these selection experiments were 5′-CGAAGCGCTAGAACAT-N_{30}-AGCTGATCCTGATGG-3′ (without the extra stem loop) and 5′-CGAAGCGCTAGAACATGCCG-N_{30}-CGGCCAGCTGATCCTGATGG-3′ (with the extra four base-pair stem loop). They were synthesized on the ABI 394 DNA/RNA oligonucleotide synthesizer. All other sequences were prepared by solid-phase synthesis at Integrated DNA Technologies (Coralville, IA). All oligonucleotides were purified by denaturing 20% or 8% PAGE.

All reactions were performed at room temperature. Centrifugations were performed for 1 min at 1000 RCF (Relative Centrifugal Force) with 5494 model Eppendorf centrifuge in Micro Bio-spin chromatography columns (Bio-Rad). Incubations of DNA with agarose beads were done inside 2.0 mL low retention microcentrifuge tubes (Fisher) on the nutator (Fisher 260100F). Scintillation counter model used was Beckman LS6500.

10 μL of underivatized agarose beads were washed with 5×250 μL of selection buffer (150 mM NaCl and 40 mM $MgCl_2$ with 50 mM Tris, pH 9.0 or 40 mM $MgCl_2$ +

20 mM $MnCl_2$ with 50 mM HEPES, pH 7.5) inside a Micro Bio-spin chromatography column (Bio-Rad) for 1 min and centrifuged. Then the filtrate was discarded. Heat-denatured (95° C for 5 min, on ice for 10 min) selection pool (1 pmol of 5′-radiolabelled pool:199 pmol non-radiolabeled pool for round 1 or radiolabeled PCR product for rounds 2+) dissolved in 400 μL of selection buffer was mixed with 20 pmol of blocking oligonucleotide 5′-$(AAC)_{20}$-3′ and the washed underivatized agarose beads inside a microcentrifuge tube and incubated at room temperature on a nutator for 30 min. The mixture was then transferred to a spin column and centrifuged. The beads were washed with 2×200 μL selection buffer and the flowthrough fractions combined. The underivatized agarose beads were mixed with water and retained for scintillation counting.

The ATP-derivatized agarose beads were obtained from Jena Biosciences, Inc (Jena, Germany), catalog number AC-101 [(ATP - Immobilized γ-Aminophenyl-ATP (C_{10}-spacer)], 20 μmol/mL ATP concentration. In every selection round, 300 nmol of immobilized ATP were washed with 10×250 μL of selection buffer, then the mixture was centrifuged and the resulting filtrate discarded. 20 pmol of the blocking oligonucleotide was mixed with the washed ATP-derivatized beads and 400 μL of selection buffer. The mixture was incubated on the nutator inside an Eppendorf tube and then transferred to the spin column and centrifuged; the resulting filtrate was discarded. The ATP-derivatized beads were then mixed with the flowthrough containing selection pool from the previous step (800 μL total volume) and incubated on a nutator for 1 h. The mixture was transferred to the spin column and centrifuged. The flowthrough was retained for scintillation counting. The ATP-derivatized beads together with bound DNA sequences inside the spin column were washed with 10×200 μL of selection buffer (following addition of the buffer

the mixture was incubated for 3 min for a total incubation time of 30 min). Flowthrough fractions from every two washes were combined and retained for scintillation counting.

To elute ATP-binding DNA sequences, 300 μL of ATP elution buffer (20 μM ATP, 150 mM NaCl and 40 mM $MgCl_2$ with 50 mM Tris, pH 9.0 or 40 mM $MgCl_2$ + 20 mM $MnCl_2$ with 50 mM HEPES, pH 7.5) was mixed with the washed beads from the previous step inside 1.7 mL Eppendorf tube and incubated for 1 h (2 h for round 1) on the nutator. After that, this mixture was transferred to the spin column and centrifuged. 1 h incubation followed by centrifugation was repeated and the two elution fractions were combined and retained for scintillation counting. The spin column was washed with 2×200 μL of selection buffer, combining the flowthrough fraction for scintillation counting. For selections in which DNA sequences were eluted with EDTA, the above ATP elution step was followed by two successive 1 h incubations of the ATP-derivatized beads in 300 μl of EDTA elution buffer (5 mM EDTA, 150 mM NaCl and 40 mM $MgCl_2$ with 50 mM Tris, pH 9.0 or 40 mM $MgCl_2$ + 20 mM $MnCl_2$ with 50 mM HEPES, pH 7.5) each followed by centrifugation and retention of the combined flowthrough fractions. The spin column was washed with 2×200 μL of selection buffer combining the flowthrough fraction for scintillation counting. The underivatized agarose beads were mixed with water and retained for scintillation counting.

DNA sequences eluted with either the ATP elution buffer or the EDTA elution buffer were PCR-amplified after removing salts by precipitating DNA with ethanol and drying the pellets in a SpeedVac vacuum concentrator. The forward primer, which was used to initiate Taq-catalyzed synthesis of a pool molecule, had the sequence 5′-CGAAGCGCTAGAACAT-3′. The reverse primer, which was used to initiate Taq-

catalyzed synthesis of the reverse complement of a pool molecule had the sequence 5′-$(AAC)_4$XCCATCAGGATCAGCT-3′ (where X denotes Glen Spacer 18, which is a PEG spacer that stops extension by Taq polymerase and leads to a size difference between the two PCR product strands). The amplification was performed in a total volume of 100 µL for 10 cycles (94 ºC for 30 s, 47 ºC for 30 s, 72 ºC for 30 s) by Taq polymerase with 1 µM of forward primer, 0.25 µM of reverse primer and 200 µM of each dATP, dCTP, dGTP and dTTP, in addition to a trace amount of γ-^{32}P-dCTP are added in 30-cycle PCR. Phenol-chloroform-isoamyl alcohol extraction was performed and ~100 µL of the aqueous PCR product solution was taken out. 1 µL aliquot of this solution was used as the template for a 30-cycle PCR in a total volume of 50 µL. The PCR product was purified by 8% PAGE. The shorter PCR product was excised, extracted, precipitated by ethanol and used for the next round of aptamer selection with ATP-derivatized agarose beads.

3.4.2 Cloning of individual sequences

The selection process led to a pool enriched with ATP-binding DNA molecules. These DNA sequences had same number of nucleotides (approximately same molecular weight) but differed in their nucleotide compositions. In order to obtain sequences of individual ATP aptamer candidates, molecular cloning was performed.

The pool synthesized in the 10-cycle PCR step of the selection round of interest was used as a PCR template to generate double-stranded DNA for cloning. This PCR product was purified on 2% agarose gel and inserted into a TOPO vector containing an ampicillin resistance gene (Invitrogen). Competent *E. coli* cells were transformed with

these plasmids. These cells were spread on a cell culture media plate containing ampicillin and the plate was incubated at 37^{o}C for 14-16 h. During this time, cells divided, forming either white colonies, in which the double-stranded DNA PCR product was inserted into the vector, or blue colonies, in which the vector did not include double-stranded PCR product. White colonies were picked from these plates and grown at 37^{o}C for 18 h in culture media containing 50 μg/mL ampicillin.

Every white colony was generated from a cell containing a plasmid that corresponded to one particular aptamer candidate sequence. During the cloning procedure, individual DNA sequences were amplified inside *E. coli*. The DNA from these cells was extracted using miniprep kits (Thermo Fisher) and the proper insert size was verified by plasmid digestion with *EcoRI*. Sequence compositions of individual aptamer candidates were provided by the Core Sequencing Facility at the University of Illinois.

3.4.3 Testing of aptamer candidate sequences by bead binding

Bead-binding assays were performed essentially as described for the aptamer selection experiments, up to and including the 1 h incubation of DNA sequences with ATP-derivatized agarose. Next, washes with 10×100 μL of selection buffer were performed, and all flowthrough fractions were retained. After that, seven 30 min incubations followed by elutions via centrifugation were done. Each elution buffer was prepared similarly to that for aptamer selections except free ATP was present in each wash at an increasing concentration (30 nM, 100 nM, 300 nM, 1 μM, 3 μM, 10 μM, 30 μM and 100 μM). Eluted solutions from each wash in addition to the agarose beads were retained for scintillation

counting. Each aptamer candidate's DNA molecule was 5′-radiolabeled with γ-^{32}P-ATP by PNK and used in trace amount together with 50 pmol of unlabeled DNA.

3.4.4 DMS probing assays

0.2 pmol of 5′-radiolabeled aptamer candidate dissolved in 9 μL selection buffer was incubated at room temperature for 30 min. An 11 μL aliquot of ATP dissolved in selection buffer was added for a final ATP concentration of 1 nM, 10 nM, 100 nM, 1 μM, 10 μM or 100 μM. A control reaction with no ATP was also performed. All samples were incubated at room temperature for 1 h. Methylation of DNA was initiated by adding 2.25 μL of freshly prepared aqueous 1.2% DMS to each sample. After 30 min incubation, DNA was precipitated with ethanol and the dried pellet was dissolved in 50 μL of 10% piperidine. The mixture was heated for 30 min at 90°C followed by cooling on ice for 30 min after which the samples were dried in the SpeedVac. Each sample was resolved on PAGE and exposed to a Phosphoimager screen.

Sequences of the oligonucleotides used in DMS probing assays shown in Fig. 3.7 were:

9D3: 5′-GCCGACCCCAGGACCAAAGTTCGACAGAGCGAGACGGCCAGCTGATC CTGATGG-3′,

DH: 5′-ACCTGGGGGAGTATTGCGGAGGAAGGT-3′,

MB: 5′-CCACTGAGGGGCGAGGAGCAGAAACTGCGGGGAGTGG-3′.

3.4.5 Equilibrium filtration assays

Each ATP aptamer candidate, DH and MB ATP aptamers as well as $(AAC)_9$ oligonucleotide reference sample (doesn't bind ATP) was dissolved in a 200 μL solution and was incubated in Amicon 3K Ultra centrifugal filters (Millipore) at room temperature for 10 min. Each mixture contained 30 μM ATP, 2.5 nM γ-^{32}P-ATP, and DNA at one of 100 nM, 200 nM, 500 nM, 1 μM, 2 μM, 5 μM, 10 μM, 20 μM and 100 μM concentrations. Additionally, three salt/buffer conditions were used:

i) 5 mM $MgCl_2$, 20 mM Tris, pH 7.6, 300 mM NaCl;

ii) 40 mM $MgCl_2$, 20 mM $MnCl_2$, 50 mM HEPES, pH 7.5, 150 mM NaCl;

iii) 40 mM $MgCl_2$, 50 mM CHES, pH 9.0, 150 mM NaCl.

After incubation, the mixtures were centrifuged in Amicon filters at 8000 RCF for 15 min resulting in 40-80 μL volume in the filtrate. 20 μL volumes of both filtrate and retentate fractions from each sample were used for scintillation counting.

3.5 References

1. Ellington, A. D.; Szostak, J. W., In vitro selection of RNA molecules that bind secific ligands. *Nature* **1990,** *346*, 818-822.
2. Tuerk, C.; Gold, L., Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA Polymerase. *Science* **1990,** *249*, 505-510.
3. Silverman, S. K., Artificial functional nucleic acids: aptamers, ribozymes and deoxyribozymes identified by in vitro selection. In *Functional Nucleic Acids for*

Analytical Applications, Y. Li and Y. Lu, eds.; Springer Science + Business Media, LLC: New York, **2009**.

4. Dieckmann, T.; Butcher, S. E.; Sassanfar, M.; Szostak, J. W.; Feigon, J., Mutant ATP-binding RNA aptamers reveal the structural basis for ligand binding. *J. Mol. Biol.* **1997,** *273*, 467-478.
5. Meyer, M.; Scheper, T.; Walter, J. G., Aptamers: versatile probes for flow cytometry. *Appl. Microbiol. Biotechnol.* **2013,** *97*, 7097-7109.
6. Amaya-Gonzalez, S.; de-los-Santos-Alvarez, N.; Miranda-Ordieres, A. J.; Lobo-Castanon, M. J., Aptamer-Based Analysis: A Promising Alternative for Food Safety Control. *Sensors* **2013,** *13*, 16292-16311.
7. Wang, C.; Liu, B.; Lu, J.; Zhang, G.; Lu, A. P., Strategies for Combination of Aptamer and Targeted Drug Delivery. *J. Nanosci. Nanotechnol.* **2014,** *14*, 501-512.
8. Zhou, W. Z.; Huang, P. J. J.; Ding, J. S.; Liu, J., Aptamer-based biosensors for biomedical diagnostics. *Analyst* **2014,** *139*, 2627-2640.
9. Sassanfar, M.; Szostak, J. W., An RNA motif that binds ATP. *Nature* **1993,** *364*, 550-553.
10. Huizenga, D. E.; Szostak, J. W., A DNA aptamer that binds adenosine and ATP. *Biochemistry* **1995,** *34*, 656-665.
11. Lin, C. H.; Patel, D. J., Structural basis of DNA folding and recognition in an AMP-DNA aptamer complex: distinct architectures but common recognition motifs for DNA and RNA aptamers complexed to AMP. *Chem. Biol.* **1997,** *4*, 817-832.
12. Barbu, M.; Stojanovic, M. N., A fresh look at adenosine-binding DNA motifs. *ChemBioChem* **2012,** *13*, 658-660.

13. Baldwin, A. S., Oettinger, M. and Struhl, K. Methylation and Uracil Interference Assays for Analysis of Protein-DNA Interactions. *Current Protocols in Molecular Biology*, 36:12.3:12.3.1–12.3.7, **2001**.
14. Kato, T.; Takemura, T.; Yano, K.; Ikebukuro, K.; Karube, I., In vitro selection of DNA aptamers which bind to cholic acid. *Biochim. Biophys. Acta.* **2000,** *1493*, 12-18.

Chapter 4: Selections of DNA Enzymes with Discrete Catalyst and ATP Aptamer Domains[a]

4.1 Introduction

Nucleic acids, peptides, sugars and other substrates attached to a random DNA pool have been used to identify deoxyribozymes by in vitro selection.[1-7] DNA-hydrolyzing deoxyribozymes, described in Chapter 2, cleave an all-DNA substrate with an external water molecule in aqueous solution. In these and similar cases, much of the catalyst-substrate binding interaction—and therefore the corresponding binding energy—is provided by the Watson-Crick base pairing. These straightforwardly preprogrammed binding contacts enable the selection process to focus upon finding initially random DNA sequences that are tasked primarily with catalysis. However, for small-molecule substrates that inherently cannot engage in extensive Watson-Crick interactions, the initially random DNA region (e.g., N_{40}) must instead simultaneously bind the substrate and catalyze the desired reaction, rather than solely participate in catalysis. DNA-hydrolyzing deoxyribozymes, assisted by metal ion cofactors, bind a water molecule (although its 55 M concentration does not require substantial binding specificity) and catalyze strand cleavage.[8] The work described in this chapter was based on an explicitly modular approach to deoxyribozyme catalysis, in which a predefined DNA aptamer (binding) domain engages in aptamer-like non-Watson-Crick contacts with the small-molecule substrate. This could allow identification of a distinct "catalytic" (enzyme) domain from a random

[a] The material described in this chapter has been published and is used here with permission: Dokukin, V.; Silverman, S. K., A Modular Tyrosine Kinase Deoxyribozyme with Discrete Aptamer and Catalyst Domains, *Chem. Commun.* **2014**, *50*, DOI: 10.1039/c4cc04253k.

DNA region through in vitro selection experiments. Emerging sequences from the initially random regions (conjugated to the DNA aptamer module) would be devoted exclusively to catalysis. Inclusion of a binding domain adjacent to the random region of a DNA pool was hypothesized to aid in finding more numerous and higher-yielding deoxyribozymes than via the regular, nonmodular, selection experiments, provided that the DNA enzymes make functional use of the aptamer module.

We chose ATP (adenosine triphosphate) as the small-molecule substrate for this work. ATP plays a defining role in biology, serving as a cofactor or a building block in numerous cellular processes essential to life. Identification of DNA catalysts capable of using ATP as substrates in reactions that are generally catalyzed by protein enzymes (such as oligonucleotide phosphorylation or adenylylation) was achieved previously.[9,10] Separately, DNA aptamers that bind ATP with K_d values in the low micromolar ranges have also been found.[11,12] One particular Mg^{2+}-dependent ATP DNA aptamer domain[11] that has been structurally characterized[13] is the basis for a wide range of experiments that involve sensors[14-29] or other applications.[30-33] Here we wished to determine whether providing a predefined ATP aptamer module can lead to deoxyribozymes that make functional use of this domain during catalysis. There is a natural nucleic acids precedent of the modularity phenomenon in that the group I intron ribozyme is modular in binding its RNA substrate.[34] Of course, many natural protein enzymes are functionally modular.[35] Additionally, we sought to assess whether the availability of the ATP aptamer module provides a more robust access by in vitro selection to new deoxyribozymes for chemical reactions that require ATP substrates, leading to more numerous or faster DNA catalysts.

4.2 Results and Discussion

4.2.1 In vitro selection of deoxyribozymes for phosphorylation and adenylylation of the 5′ DNA terminus

Our initial efforts focused on identification of DNA catalysts capable of transferring either a phosphoryl or an adenylyl functional group to, respectively, the 5′-hydroxyl or the 5′-phosphate termini of single-stranded DNA. ATP serves as a natural substrate for both of these reactions (Fig. 4.1). Phosphorylation of the 5′-OH terminus of DNA is catalyzed by

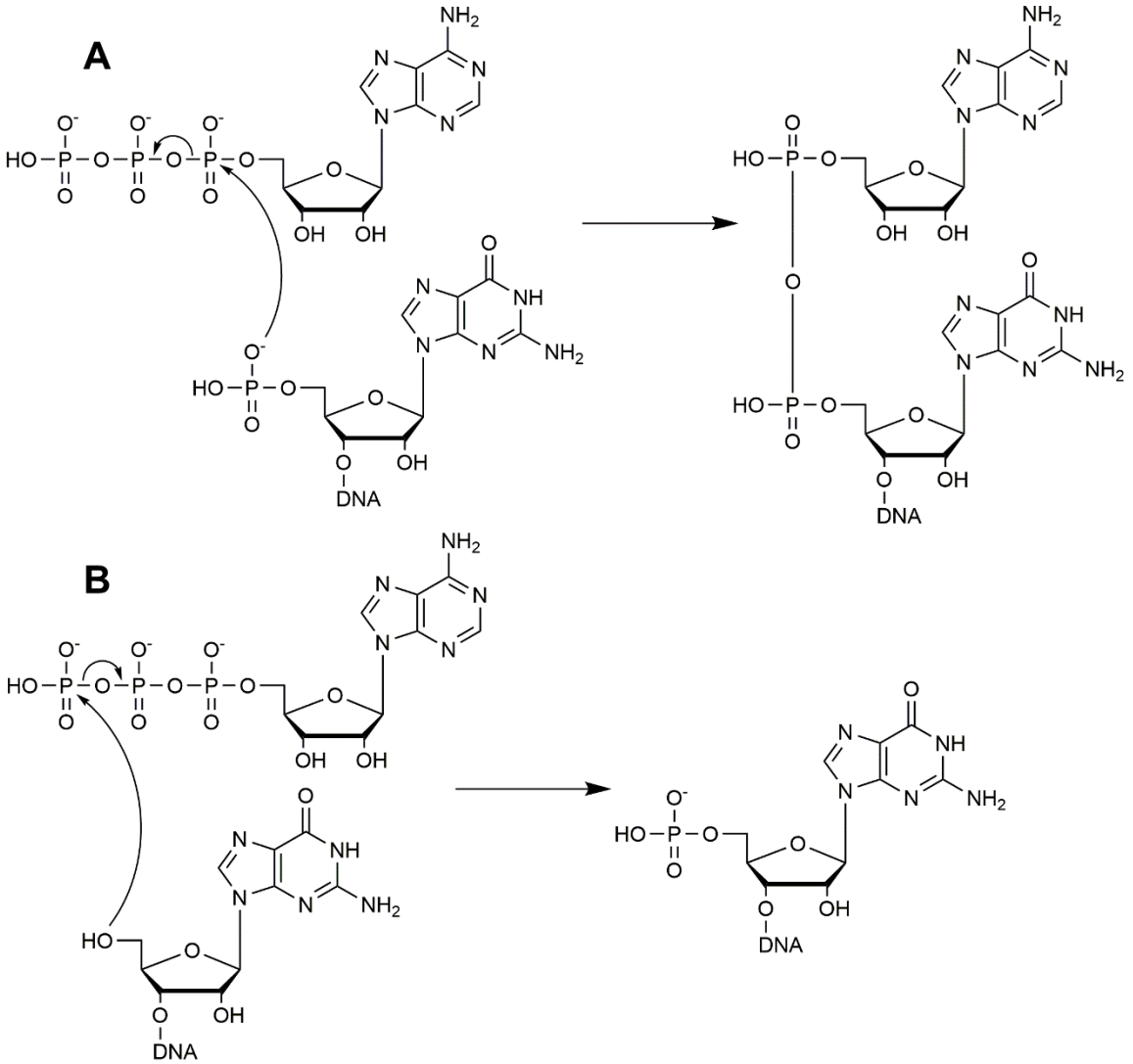

Figure 4.1: The two reactivities with ATP for which deoxyribozymes were selected via the modular approach: adenylylation (A) and phosphorylation (B).

polynucleotide kinases, which are also extensively used in biochemical laboratories that study nucleic acids. Nonmodular polynucleotide kinase deoxyribozymes have been identified by others.[10] To find modular DNA catalysts of this reaction, we started four selection experiments with an integrated DNA aptamer domain.

On the other hand, adenylylation of a 5′-terminal DNA phosphate is the first step of the ligation mechanism in which two DNA strands are joined together. This reaction is catalyzed in nature by protein ligases. Ligation is an essential laboratory procedure for the molecular cloning of DNA. Deoxyribozyme-catalyzed adenylylation of DNA was also achieved previously.[9] In that work, catalysts were identified via selection experiments from N_{70} random-sequence pools in the absence of an ATP aptamer module. Encouraged by these results, we initiated three selection experiments using the modular approach to identify DNA-adenylylating DNA enzymes with distinct catalytic and binding domains.

Because neither 5′-terminal phosphorylation nor adenylylation significantly increases charge or mass of a pool molecule linked to the nucleic acid substrate, catalytically active sequences cannot be readily separated from inactive ones on PAGE. A strategy to capture active sequences was previously developed.[9,10] Splint ligation reactions (also referred to as the capture reactions) of the 5′-modified DNA molecules with acceptor oligonucleotides were employed to achieve separation as the ligation products had higher mass and charge than the reactants. T4 DNA ligase enzyme was used to catalyze these reactions. To avoid propagation of aberrantly migrating sequences, single-stranded DNA capture oligonucleotides of alternating lengths were used in successive selection rounds.

5′-Phosphorylation of the donor DNA sequence is required for ligation by the DNA ligase enzyme in the presence of ATP. A splint molecule complementary to the ligated DNA molecules is also required for this reaction. Only those DNA substrates that acquired the 5′-phosphate modification could be ligated, therefore enabling the separation of deoxyribozyme kinases from the rest of the random-sequence DNA pool molecules (Fig. 4.2A). For adenylylation activity, the splint ligation reactions were performed in absence of ATP. The first step of a DNA ligation reaction catalyzed by the DNA ligase enzyme is the transfer of the adenylyl group (for which ATP is the donor) to the 5′-phosphate of DNA. This reaction intermediate readily undergoes the second step of the ligation reaction and becomes covalently linked to the 3′-end of the acceptor DNA sequence. Because during the selection step, the substrate could be preadenylylated (transfer of an adenylyl group to the 5′-phosphate) by catalytically active random-sequence DNA pool members, such sequences would be ligated (captured) to the acceptor oligonucleotide without the otherwise-required ATP (Fig. 4.2B). However, upon initiation of these selection experiments, we realized that the DNA ligase enzyme used for the capture step of the selection experiments is provided in the preadenylylated (activated) form in the stock tube. This resulted in the capture of sequences that did not undergo the required adenylylation, as manifested by observation of ligation products co-migrating with the oligonucleotide standards on PAGE in a very early (Round 2) selection rounds. When starting with a completely random pool, very rare catalytically active sequences are not expected to begin saturating the pool at such an early point. This problem was not described in the report of successful discovery of DNA-catalyzed DNA adenylylation,[9] perhaps due to the different vendors from which T4 DNA ligase was purchased (New

England Biolabs, from the previous report vs. Fermentas, here). To deactivate the preadenylylated DNA ligase molecules, they were incubated with an excess of 5′-phosphorylated DNA sequences prior to their addition to the pool to catalyze the capture step. This strategy was successful, as captured pool sequences were not observed during early rounds after the selection was reinitiated.

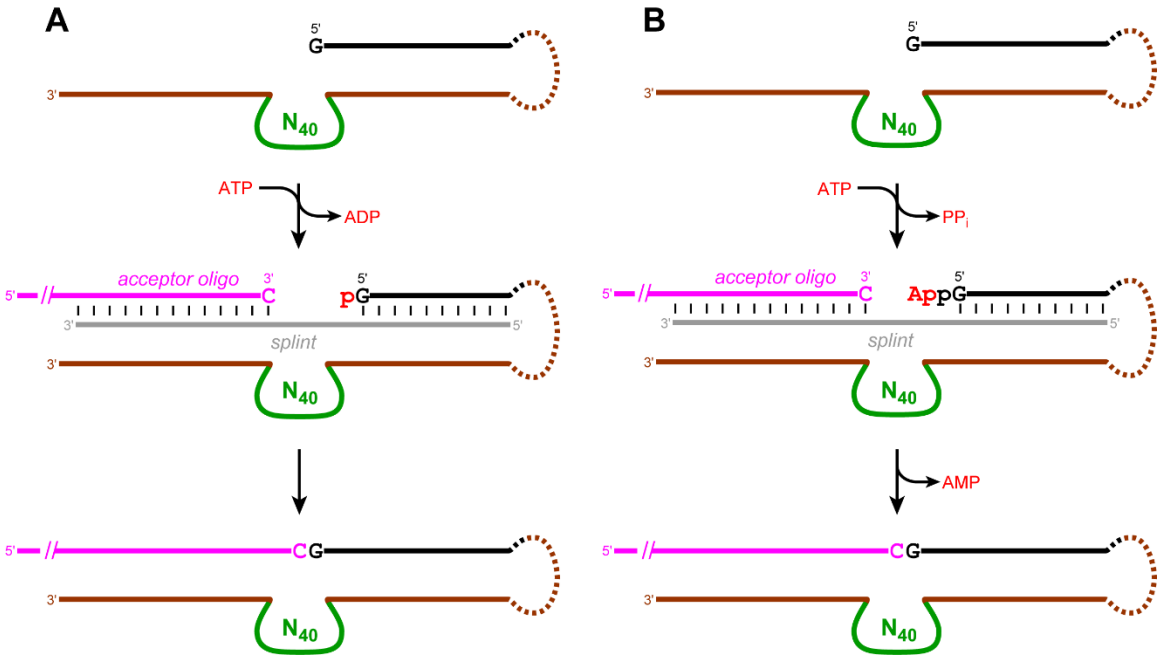

Figure 4.2: Selections of deoxyribozymes for DNA phosphorylation (A) and DNA adenylylation (B). Only the pool without the aptamer region is shown. The acceptor (capture) oligonucleotide is provided after the selection step. The capture step is performed in the presence of ATP in phosphorylation selections and in the absence of ATP in adenylylation selections.

Three N_{40} random-sequence DNA pools were synthesized, one of which did not include an aptamer module while each of the other two had either the 27-nucleothide DH (reported in 1995 by **D**avid **H**uizenga and Jack Szostak[11]) or the 37-nucleotide MB (reported in 2012 by **M**ihaela **B**arbu and Milan Stojanovic[12]) DNA aptamer for ATP. For two of the three pools, the aptamer modules were placed immediately at the 3′-end of the

random region, flanked on the other side by the invariant primer-binding sequence. The third pool lacked the aptamer domain. The set of two primer-binding regions was identical for all three pools. The random pool that did not include an ATP aptamer was 73 nucleotides long while modular incorporation of DH and MB aptamer sequences resulted in enlargement of the respective pools to 100 and 110 nucleotides. Due to different pool lengths, three sets of oligonucleotide standards (to be resolved on PAGE with selection pools) were used after each step of the selection process. A total of seven selection experiments were initiated and for all, the key selection reaction was incubated with 1 mM ATP for 14 h at 37 °C. With four selection experiments we were seeking to isolate polynucleotide kinase deoxyribozymes [one with either MB or no ATP aptamer module and two with DH (Mg^{2+} + Mn^{2+} and Mg^{2+} alone) ATP aptamer module-containing random pools], while three selections were aimed at identification of DNA-adenylylating DNA catalysts (one with each DH, MB and no ATP aptamer module-containing random pool). In addition, the incubation mixture included 40 mM $MgCl_2$ and 20 mM $MnCl_2$ with 50 mM HEPES, pH 7.5 and 150 mM NaCl (six of seven selections) or 40 mM $MgCl_2$ with 50 mM CHES, pH 9.0 and 150 mM NaCl (one of seven selections). Although the DH aptamer was identified through incubations at room temperature, extensive studies indicate that its ATP binding is at least as strong at the elevated temperature of 37 °C as at room temperature.[36,37]

Each selection was iterated for eight to twelve rounds. All seven selections resulted in substantial capture yields (up to 23%). This yield value should be compared to the maximum capture yield of prephosphorylated or preadenylylated substrates of 60-70%. Prior to cloning individual sequences, the entire pool for each selection round that displayed activity was tested intermolecularly (in trans), where substrate molecules were

not covalently attached to pool sequences, although extensive Watson-Crick hydrogen-bonding still allows strong interactions. Notably, a single-stranded DNA loop is formed when pool and substrate sequences are connected during selection rounds. Unfortunately, when the pools were tested intermolecularly, neither phosphorylation nor adenylylation of substrates was observed. This result indicated that none of the pools had sequences capable of supporting catalysis intermolecularly and the absence of the single-stranded loop explains this outcome. However, these results (especially for selections with the pool lacking an aptamer module) are in disagreement with previous reports of successful identification of deoxyribozymes with 5′-phosphorylation as well as 5′-adenylylation activities.[9,10] This outcome means that DNA catalysts of DNA phosphorylation or adenylylation were inaccessible via the selection process as outlined here. In order to attempt to obtain DNA enzymes with these activities, the conditions of the selection step need to be changed. Instead, we set aside such efforts and focused on a different activity.

4.2.2 In vitro selection of deoxyribozymes for phosphorylation of a tyrosine hydroxyl

Guided by recent discovery of tyrosine deoxyribozyme kinases in the Silverman laboratory[38] and by an overall greater scientific interest in protein (rather than nucleic acid) modifications, we shifted our focus towards identification of modular deoxyribozymes capable of supporting amino acid phosphorylation. The recently reported tyrosine kinase deoxyribozymes emerged from selections in which GTP (guanosine triphosphate) was used as the phosphoryl donor. Upon further study of these DNA enzymes, ATP was also found

to support catalytic activity for one of them. The laboratory's earlier report of deoxyribozymes for tyrosine dephosphorylation indicated additional competence of DNA to support catalysis involving a protein side chain and phosphoryl transfer.[39] Having achieved tyrosine kinase reactivity (by nonmodular DNA catalysts) in the Silverman laboratory and possessing the knowledge and strategy for identifying such deoxyribozymes, we decided to evaluate our ability to find DNA kinases via the modular approach. We expected an improved access to catalysts via this approach, as described above.

4.2.3 Tyrosine kinase selections with 1 mM ATP

Six DNA pools were synthesized with one of two ATP aptamers (DH and MB) adjacent to the N_{30}, N_{40} and N_{50} DNA random regions (Table 4.1). Here, the aptamer sequences also served as primer-binding regions for PCR. Three additional pools lacking a predefined aptamer domain were prepared. Instead of the aptamer module, a 16-nucleotide primer-binding sequence was located adjacent to the N_{30}, N_{40} and N_{50} DNA random regions. The length of the random region was previously shown to be a critical experimental variable.[40] Selection experiments with these nine pools were performed for tyrosine kinase activity towards a DNA-anchored CAAYAA hexapeptide substrate, which included a tyrosine acceptor of the phosphoryl group while ATP served as the phosphoryl donor. The hexapeptide was connected to the DNA anchor by a hexa(ethylene glycol), or HEG, tether (Fig. 4.3).

Name	**FD1**	**FE1**	**FF1**	**FG1**	**FH1**	**FJ1**	**FK1**	**FL1**	**FM1**
Random region	N_{30}	N_{40}	N_{50}	N_{30}	N_{40}	N_{50}	N_{30}	N_{40}	N_{50}
ATP aptamer	no aptamer			DH aptamer			MB aptamer		

Table 4.1: Alphanumeric designation of each selection experiment with its corresponding DNA random region length and the ATP DNA aptamer (if any) integrated. For the FD1-FF1 selections, a 16-nucleotide long fixed primer-binding region was used in place of an aptamer module for PCR amplification purposes.

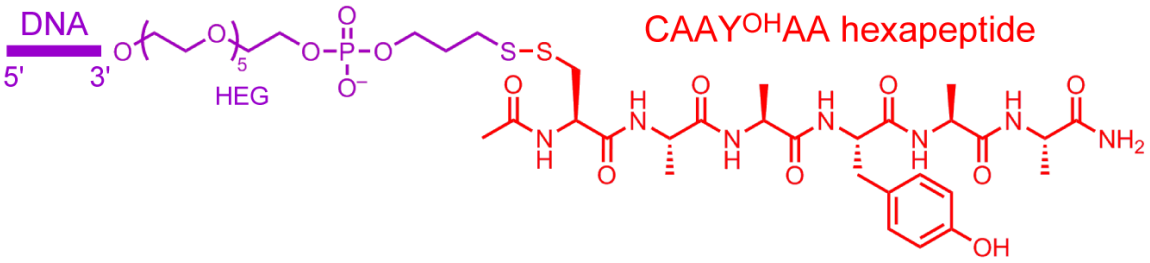

Figure 4.3: Structure of the DNA-anchored CAAYAA hexapeptide phosphorylation substrate. Note the hexa(ethylene glycol), or HEG, tether connecting the DNA oligonucleotide anchor (purple bar) to the peptide.

The incubations during the key selection step were performed with 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl and 1 mM ATP at 37 °C for 14 h. Because the K_d for the isolated DH ATP aptamer is on the order of 6 μM[11] and the K_d for the isolated MB aptamer is in the similar range,[12] 1 mM ATP concentration was expected to be fully saturating for each aptamer domain.

Similar to the previous selection effort aimed at identification of polynucleotide kinases, here phosphorylation of a tyrosine side chain does not alone enable the selection process because phosphorylated substrate-pool conjugates cannot be readily separated from unreacted ones by PAGE. To enable the selection process, we exploited a previously identified deoxyribozyme, 8VP1, that is capable of attaching an RNA strand specifically to phosphorylated tyrosine, Y^P, within a peptide substrate while discriminating strongly against the unphosphorylated analogue, Y^{OH} (Fig. 4.4).[41] This deoxyribozyme was used earlier to help identify nonmodular tyrosine kinase DNA catalysts.[38] 8VP1 provides ca. 25–

35% capture yield, which sets a corresponding upper limit on the observable yield from each random pool. In this work, 8VP1 is referred to as the "capture" deoxyribozyme because during the selection process it allows to capture (by PAGE shift) specific DNA sequences that phosphorylate tyrosine. 8VP1 plays the role of T4 DNA ligase used to capture catalytic sequences in the previous set of selection experiments.

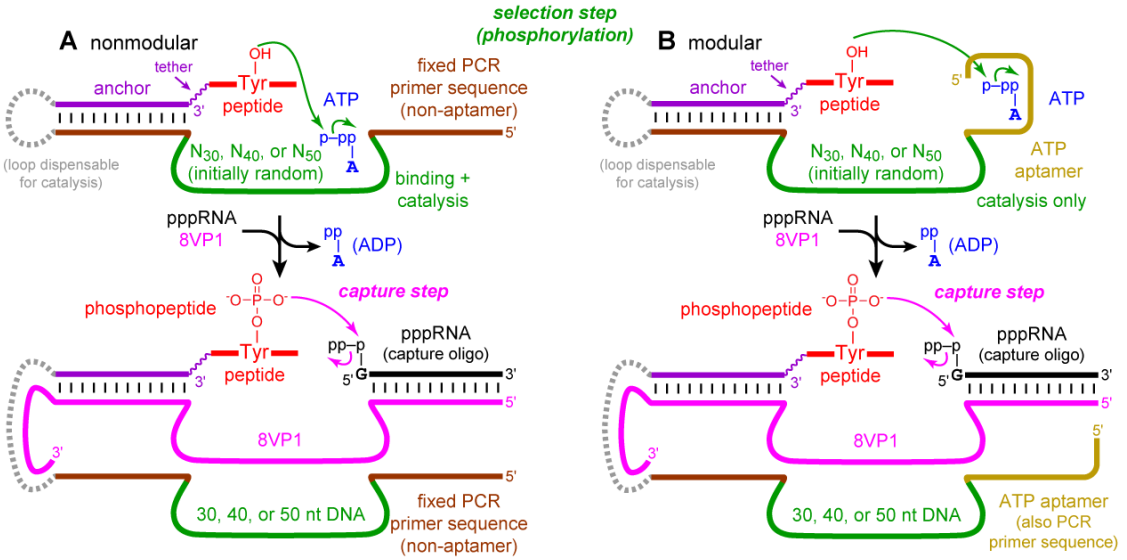

Figure 4.4: Selection of nonmodular (A) and modular (B) tyrosine kinase deoxyribozymes. DNA-catalyzed peptide phosphorylation is followed by "capture" using the phosphotyrosine-specific 8VP1 deoxyribozyme and a 5′-triphosphorylated RNA oligonucleotide, enabling PAGE-shift separation of kinase deoxyribozymes. The capture step is followed by PCR (not shown) to initiate the next selection round.

After iterating the selections for 14 rounds, only the FD (N_{30}, no ATP aptamer module) and FG (N_{30}, DH ATP aptamer module) selections displayed up to 14% and 9% catalytic activity, respectively. These yield values should be compared to the ca. 25–35% maximum capture yield by 8VP1, as indicated above. These yields were observed in rounds 10 through 13 and individual sequences form each of the two pools at round 13 were cloned in and tested for activity. Although many clones were catalytically active, the highest phosphorylation yields were under 10%. Upon sequencing, all of the active 13FD1 clones

were found to be either single mutants or exactly identical to 8EA101, a nonmodular ATP-dependent tyrosine kinase deoxyribozyme previously identified in the Silverman laboratory.[38] Apparently, the 8EA101 sequence contaminated the FD1 selection, which was possible because the primer-binding sequences in both the EA1 and the FD1 selections were identical. This outcome was not surprising, noting previous instances of sequence contamination propagated by PCR. However, after cloning and sequencing of the 13FG1 selection pool (N_{30}, DH ATP aptamer module), one unique sequence, 13FG108, was identified.

When prepared by solid-phase synthesis, the 13FG108 deoxyribozyme provided up to 30% tyrosine phosphorylation yield after 16 h. However, 13FG108 deoxyribozyme variant lacking the 27-nucleotide DH aptamer module catalyzed tyrosine phosphorylation with yields comparable to those for the full-length 13FG108. The dispensable nature of the aptamer module indicated that this domain provided no functional contribution to catalysis by 13FG108 DNA enzyme. Further evidence, based on mutational analysis, supported this conclusion. For this work, eight variants of the 13FG108 DNA catalyst with strategic single-nucleotide mutations in the aptamer sequence were synthesized. The study in which the DH ATP aptamer domain was initially identified also evaluated the importance of many individual nucleotides to ATP binding by the isolated aptamer.[11] Five of the synthesized 13FG108 variants included mutations that were shown to inhibit binding of ATP while three were known to be innocuous. When assayed with the hexapeptide substrate and ATP, three of the eight 13FG108 mutants (two innocuous and one deleterious) did not provide the expected outcome, therefore indicating no functional contribution of the aptamer module to catalysis. When a previously reported[38] tyrosine kinase deoxyribozyme

identified via direct (nonmodular) selections was evaluated with various concentrations of NTPs (including ATP), strong inhibition of activity above 100 μM ATP was observed. Although they were selected with 1 mM NTP, these DNA catalysts were inactive at this ATP concentration. Therefore, our inability to find truly modular deoxyribozymes for tyrosine phosphorylation in this particular selection effort was attributed to catalytic suppression induced by high concentration of ATP used in the selection process.

4.2.4 Tyrosine kinase selections with 30 μM ATP

We hypothesized that tyrosine kinase deoxyribozymes with functional aptamer modules could still be found via selection experiments if the ATP concentration during the key incubation step was reduced. 30 μM ATP (a 33-fold decrease from 1 mM ATP) was still expected to provide saturation of the ATP aptamer ($K_d \sim 6$ μM) while avoiding catalytic suppression observed in earlier selection experiments. Similar to the previous efforts, but excluding the incorporation of the 37-nucleotide MB aptamer, six selections were initiated (Table 4.2). Importantly, to avoid contamination by earlier-identified DNA catalysts, primer binding regions for all pools were now substituted with sequences that were not used previously in any tyrosine kinase selection experiments. Incubations during the key selection step for all six pools were performed in the presence of 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl and 30 μM ATP at 37 °C for 14 h.

Name	**JN1**	**JP1**	**JQ1**	**JR1**	**JS1**	**JT1**
Random region	N_{30}	N_{40}	N_{50}	N_{30}	N_{40}	N_{50}
ATP aptamer	no aptamer			DH aptamer		

Table 4.2: Alphanumeric designation of each selection experiment with its corresponding DNA random region length as well as presence or absence of the DH aptamer. For the JN1-JQ1 selections, a 16-nucleotide long fixed primer-binding region was used in place of the aptamer module.

After 13 rounds, the N_{40} selection with the ATP aptamer module (JS1) showed 0.9% yield, which rose to 3.1% at round 14 and 6.9% at round 17 (Fig. 4.5). All of these yield values should be compared to the 25–35% maximum capture yield by the 8VP1 deoxyribozyme as indicated above. In contrast, none of the selections without the aptamer module (JN1-JQ1) as well as the N_{30} (JR1) and the N_{50} (JT1) selections with the integrated ATP binding region gave more than 0.7% yield through round 18. The round 14 N_{40} modular pool (JS1) was cloned and sequenced, revealing that the population had converged on a single active sequence named 14JS101 (see Materials and Methods section for its nucleotide details). The intact ATP aptamer domain was necessarily present as part of the 14JS101 sequence, as enforced by the method of selection (the forward PCR primer used to form the deoxyribozyme pool at the end of each selection round is the ATP aptamer sequence, Fig. 4.4).

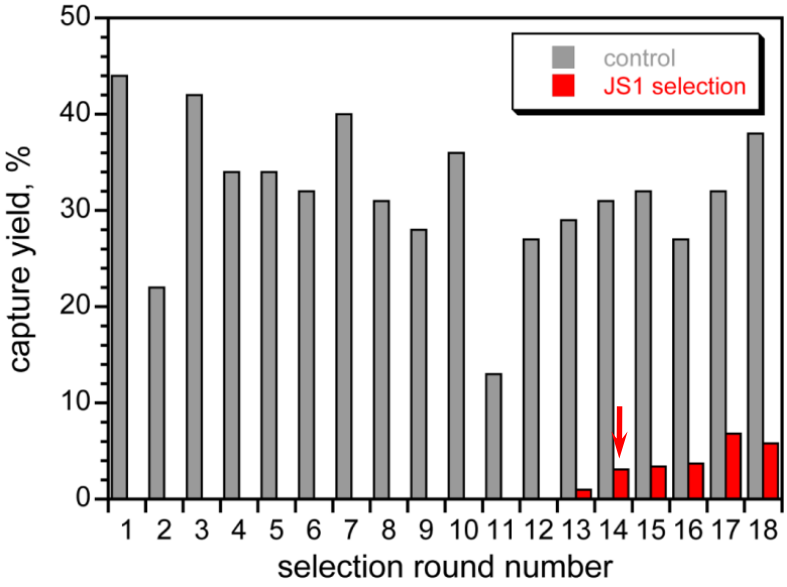

Figure 4.5: Progression of the JS1 selection experiment. The arrow marks the cloned round, number 14. Control reaction was performed in every selection round with the DNA-anchored hexapeptide with Tyr^P catalyzed by the 8VP1 deoxyribozyme.

The 14JS101 deoxyribozyme provided up to 28% phosphorylation yield in 16 h when evaluated under the same incubation conditions as were used during the selection process, in a single-turnover assay with the DNA-anchored CAAYAA substrate provided separately from the deoxyribozyme. This assay directly monitored conversion of Y^{OH} to Y^P; no capture step was involved, as addition of a negatively charged phosphate increased electrophoretic mobility of the product relative to the unreacted substrate so the two corresponding PAGE bands could be reasonably well resolved and their intensities quantified. The identity of the DNA-anchored hexapeptide phosphorylation product was validated by MALDI mass spectrometry. The phosphorylation yield of 14JS101 at 24 h was optimized to 52–64% by incubation in 40 mM Tris, pH 7.5, 1 mM $ZnCl_2$, 10 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl and 30 µM ATP at 37 °C. These optimized conditions were adopted for all further assays. The divalent metal ion dependence of 14JS101 was assayed by omitting either one or two divalent metal ions that were present

during the selection process (Fig. 4.6A). Omitting either Mn^{2+} or Mg^{2+} led to a 36-fold and eight-fold decrease in yield, respectively, while Zn^{2+} was found to be strictly required for 14JS101 catalysis. Additionally, Zn^{2+} concentration dependence of 14JS101 was also evaluated by performing reactions with a wide range of Zn^{2+} concentrations, from 0 mM to 4 mM (in 0.1 mM increments) as well as at 4.5 mM, 5.0 mM and 6.0 mM Zn^{2+} (Fig. 4.6B). The optimum yield for the reaction catalyzed by 14JS101 was achieved with 1.0 mM Zn^{2+}.

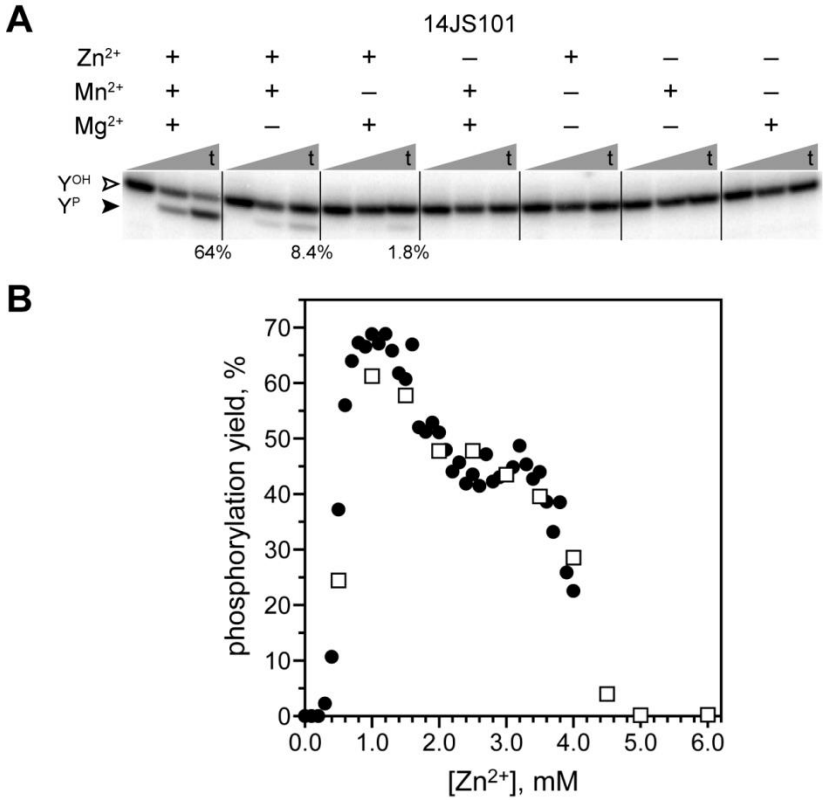

Figure 4.6: Divalent metal ion dependence of the 14JS101 deoxyribozyme. (A) Ascertaining the requirements for inclusion of Zn^{2+}, Mn^{2+} and Mg^{2+} (t = 30 s, 5 h, 24 h). (B) Determining the optimum Zn^{2+} concentration (yield values at 24 h are plotted). For all samples, the Tris concentration was held constant at 40 mM. Different symbols correspond to different data sets. The origin of the interesting and rather unusually shaped [Zn^{2+}] dependence curve is unknown.

One of the goals of this work was the identification of a truly modular deoxyribozyme, i.e., one in which an aptamer domain contributes functionally to its catalysis. To establish this for the 14JS101 DNA enzyme (as described for 13FG108 analysis), we made use of the previously published information on the importance of individual nucleotides within the DH ATP aptamer to its binding of ATP.[11] From those data, we chose five mutations anticipated to be deleterious for ATP binding and three mutations that should be innocuous, if the aptamer domain of 14JS101 contributes to catalysis by binding ATP (Table 4.3). Each of these eight mutations was made separately in 14JS101 and the kinase activity of each mutant was determined. An additional mutant of 14JS101 in which the entire 27-nucleotide aptamer domain sequence was scrambled was also examined (Fig. 4.7). For all but one mutant, the activity (or lack thereof) supported the conclusion that the ATP aptamer domain contributes directly to 14JS101 catalysis. For the A20C mutant, no activity was anticipated on the basis of the prior report, yet we observed substantial 14JS101 catalysis (~25% after 24 h). Noting that the A20G mutation of the same adenosine nucleotide is both expected and (here) observed to allow substantial activity, we surmise that the A20C mutation is simply not as deleterious in the 14JS101 deoxyribozyme context as in the isolated ATP aptamer. Deleting or replacing the ATP aptamer module of 14JS101 with the 37-nucleotide MB aptamer that binds only a single molecule of ATP (versus two ATP molecules for the parent aptamer) resulted in no 14JS101 activity under the same assay conditions (<0.5%, data not shown).

Parent		ACCTGGGGGAGTATTGCGGAGGAAGGT
Innocuous mutants	G8A	ACCTGGG**A**GAGTATTGCGGAGGAAGGT
	G9A	ACCTGGGG**A**AGTATTGCGGAGGAAGGT
	A20C	ACCTGGGGGAGTATTGCGG**C**GGAAGGT
	G21A	ACCTGGGGGAGTATTGCGGA**A**GAAGGT
	G22A	ACCTGGGGGAGTATTGCGGAG**A**AAGGT
Deleterious mutants	G5A	ACCT**A**GGGGAGTATTGCGGAGGAAGGT
	G7A	ACCTGG**A**GGAGTATTGCGGAGGAAGGT
	A20G	ACCTGGGGGAGTATTGCGG**G**GGAAGGT
Scrambled		*GTAGTTAGCGAGGCGTGACGAGGGTAG*

Table 4.3: List of the parent, eight single-mutant and one fully scrambled ATP aptamer sequences shown in the 5′ to 3′ direction. Mutated nucleotides are in boldface. The fully scramble aptamer sequence is italicized. Each aptamer variant at its 3′-end is connected to the non-aptameric portion of 14JS101. Each variant 14JS101 sequence was synthesized separately.

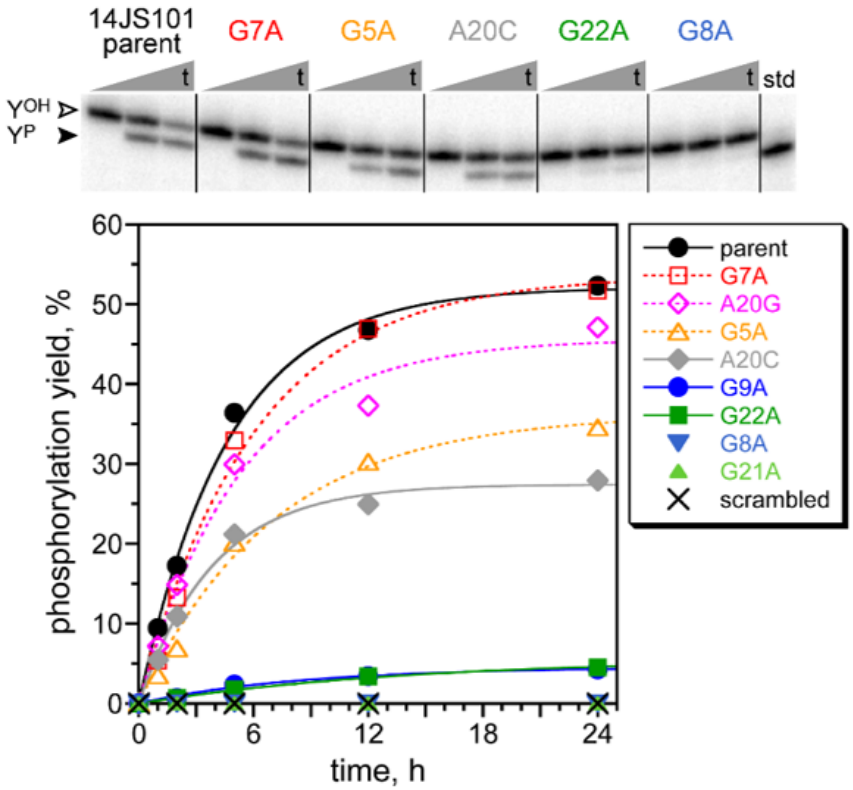

Figure 4.7: Kinetic assays of the 14JS101 parent sequence and several mutants. For the gel image, t = 30 s, 5 h, 24 h (40 mM Tris, pH 7.5, 1 mM $ZnCl_2$, 10 mM $MnCl_2$, 40 mM $MgCl_2$, 150 mM NaCl and 30 μM ATP at 37 °C). In the plot, open symbols denote mutants expected to retain activity (innocuous mutations), and filled symbols denote mutants expected to lose activity (deleterious mutants).

A first-order fit to the full kinetic plot of the parent 14S101 deoxyribozyme under optimized conditions produced a value for k_{obs} (rate constant) of 0.21 h^{-1}. However, the curve fit heavily relies on yield values at later time points, when product formation is decelerating. To obtain more reliable values, initial-rate kinetics assays were performed. When these assays are done with varying ATP concentrations, the Michaelis constant, K_m(ATP), for each active 14JS101 variant can be determined by plotting the calculated initial rate values against the corresponding concentrations of ATP. Then, each data set was fit to the following equation:

$$k_{\mathrm{obs}} = k_{\mathrm{cat}} \cdot \frac{C^n}{K_{\mathrm{m}}^n + C^n} \cdot \left(1 - \frac{C^m}{K_{\mathrm{i}}^m + C^m}\right)$$

where k_{cat} is the apparent k_{cat} value (i.e., k_{obs} at saturating ATP concentrations), C is the ATP concentration, K_m is the apparent K_m(ATP) value for productive ATP binding, K_i is the inhibition constant for unproductive ATP binding while n and m are Hill coefficients. The K_m(ATP) of 14JS101 was determined to be ca. 10 µM (Fig. 4.8). This value is in accord with the 6 µM K_d value for the isolated ATP aptamer, consistent with direct contribution of the aptamer domain to 14JS101 catalysis. The K_m(ATP) was also determined for each of the active mutants. In each case, the K_m value was 11–32 µM, suggesting at most a minor perturbation in the ability to bind ATP.

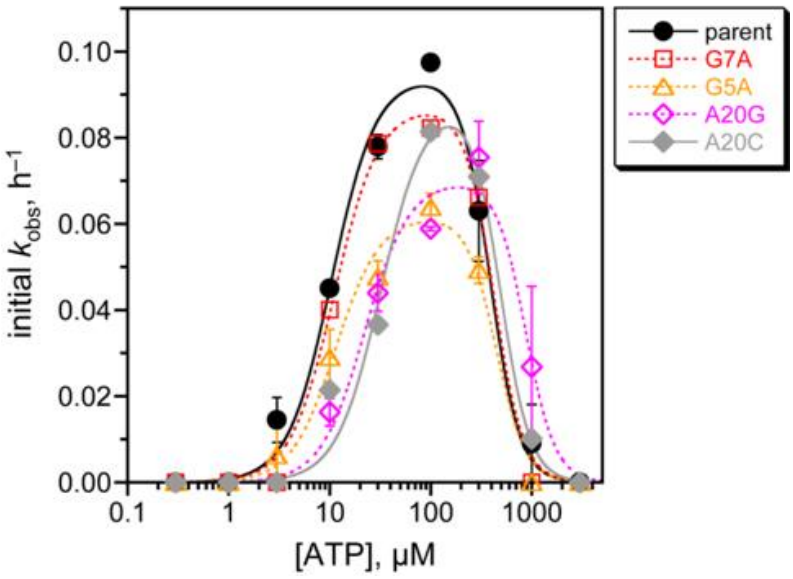

Figure 4.8: Assessment of K_m(ATP) for the parent 14JS101 sequence and four mutants using initial-rate kinetics (0–2 h). Error bars (half of range) are shown for parent, G5A and A20G (each n = 2); some of the error bars are smaller than the sizes of the data points. For G7A and A20C, n = 1. Curve were fit using Hill coefficients of 2 for activation and 3 for inhibition. The fit K_m(ATP) values are (top to bottom in legend, µM) 10.5 ± 1.2, 11.1 ± 1.2, 11.4 ± 1.6, 22.2 ± 3.6 and 32.0 ± 5.3.

The study in which the DH ATP aptamer domain was initially identified also evaluated its target specificity. In that report, none of the other three natural ribonucleotides were shown to bind the aptamer. In contrast, when ATP was substituted with dATP (which lacks the 2′-OH of ATP), the DH aptamer successfully bound this variant of the target molecule. To further verify that the aptamer domain of 14JS101 is functionally responsible for binding ATP, the DNA catalyst was tested with GTP, CTP, UTP and dATP (Fig. 4.9). All of these nucleotides have structural similarities where triphosphate functional groups (which can donate the phosphoryl group from its γ-phosphate to phosphorylate tyrosine) are connected to pentose sugars with aromatic nucleobases. Notably, nonmodular tyrosine kinase deoxyribozymes were identified with GTP as the phosphoryl donor in the Silverman lab previously.[38] When 14JS101 was tested with any of GTP, CTP or UTP, no activity was observed. However, dATP could successfully replace ATP. These findings are consistent with functional involvement of the DH ATP aptamer domain as intended.

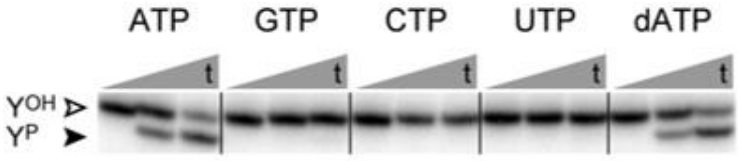

Figure 4.9: Assays of the parent 14JS101 with all four standard NTPs and dATP. t = 30 s, 5 h and 24 h.

4.3 Summary

From these findings, we conclude that 14JS101 is a truly modular tyrosine kinase deoxyribozyme, in which the ATP aptamer domain binds the small-molecule ATP substrate while the separate, initially random (N_{40}) region is responsible for catalysis. An

early ribozyme selection experiment attempted to include a predefined ATP aptamer domain to achieve polynucleotide kinase activity. However, the resulting aptamer domain became mutated during the selection process, and most of the corresponding ribozymes were unlikely to make functional use of the aptamer domain.[42,43] Similarly, selection for a self-alkylating ribozyme using a biotinylated small-molecule substrate and a predefined biotin aptamer domain led to ribozymes that do not appear to use the (mutated) aptamer.[44] In contrast to these older results, the present work establishes unequivocally that nucleic acid enzymes can use discrete aptamer and catalyst domains that together enable catalytic function. The mechanistic basis for the collaborative operation of the aptamer and catalyst domains cannot be determined from the present data. Such insights likely require high-resolution structure information, which is currently unavailable for any deoxyribozyme.[45] The DH aptamer used to identify 14JS101 binds two molecules of ATP. The available data do not establish conclusively whether or not 14JS101 requires binding of two ATP molecules. We speculate that if 14JS101 binds two ATP molecules, then only one of these molecules is properly oriented to serve as the phosphoryl donor.

A parallel goal of the present work was to evaluate whether including a predefined aptamer module leads to more robust identification of DNA catalysts. Although these experiments did successfully lead to 14JS101 as a single, explicitly modular kinase deoxyribozyme, a broader range of modular DNA catalysts was not found. Indeed, we found no deoxyribozymes at all from the analogous N_{30} and N_{50} modular selections as well as for all three N_{30}, N_{40} and N_{50} selections without the integrated aptamer region. However, previous efforts in the Silverman laboratory with N_{30} and N_{50} random regions (but no aptamer domain) and using GTP as the phosphoryl donor did provide DNA catalysts.[38]

Moreover, the k_{obs} of 0.2 h^{-1} for 14JS101 is the same as the k_{obs} of 0.2 h^{-1} for the fastest NTP-dependent deoxyribozyme in the previous effort.[38] Therefore, including a predefined aptamer module, while successful specifically for finding the 14JS101 deoxyribozyme, does not appear to be an especially helpful principle on which to design DNA catalyst selection strategies, at least for the kinase activity examined here. This conclusion is meaningful in the context of ongoing efforts to identify new deoxyribozymes that function with a wide range of small-molecule substrates. These findings establish experimentally that explicitly modular deoxyribozymes are possible, and for some purposes, pursuing such modularity may be advantageous.

4.4 Materials and Methods

4.4.1 Oligonucleotides, peptides and DNA-anchored peptide conjugates

DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) or prepared by solid-phase synthesis on an ABI 394 instrument using reagents from Glen Research (Sterling, VA). 5′-Triphosphorylated RNA (pppRNA) oligonucleotides were prepared by in vitro transcription using synthetic DNA templates and T7 RNA polymerase. All oligonucleotides were purified by 7 M urea 8%, 12% or 20% (as appropriate) denaturing PAGE with running buffer 1× TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3).

Peptides were prepared by solid-phase synthesis using Fmoc Rink amide MBHA resin, with *N,N,N′,N′*-tetramethyl-*O*-(7-azabenzotriazol-1-yl)uronium hexafluoro-phosphate (HATU) as coupling agent. Each synthesis was performed at 0.2 mmol scale,

initiated using 308 mg of Rink amide resin with a loading capacity of 0.65 mmol/g (AAPPTec). All steps were monitored by ninhydrin test. For each coupling, 5 equivalents (1.0 mmol) of Fmoc-amino acid, 4.9 equivalents (373 mg, 0.98 mmol) of HATU and 10 equivalents (350 μL, 2.0 mmol) of *N*,*N*-diisopropylethylamine (DIPEA) were mixed in 10 mL of anhydrous DMF. The coupling reaction was initiated by adding this mixture to the resin and agitating by bubbling with nitrogen for 45 min, followed by washing with DMF (3 × 10 mL). The N-terminus of the newly installed amino acid was deprotected by agitating the resin in 20% piperidine in DMF (5 mL) under nitrogen for 5 min a total of three times, each time washing with DMF (3 × 5 mL). Peptides with N-terminal cysteine were capped via acetylation of their N-terminus by agitating the resin in 50 equivalents (924 μL, 10 mmol) of acetic anhydride and 25 equivalents of DIPEA (870 μL, 5 mmol) in 5 mL of DMF under nitrogen for 30 min. The peptide was cleaved from the solid support by stirring the resin in a separate vial with a solution containing 5 mL of trifluoroacetic acid (TFA), 125 μL of water, 125 μL of ethanedithiol and 50 μL of triisopropylsilane for 90 min. The liquid solution was separated from the resin by filtration. This solution was dried on a rotary evaporator to yield a solid. To this material 20 mL of cold diethyl ether was added, and the peptide was obtained as a white solid.

For the tyrosine kinase selections, the DNA anchor oligonucleotides were as follows:

5′-GGATAATACGACTCACTAT-HEG-p-C_3-SS-C_3-OH-3′ (FD1-FM1),

5′-GGACTATTGAAAGACATAT-HEG-p-C_3-SS-C_3-OH-3′ (JN1, JQ1, JS1),

5′-GGATCAGGTTACTAATTAT-HEG-p-C_3-SS-C_3-OH-3′ (JP1, JR1, JT1)

(see first structure in Figure 4.10), where the 3′-disulfide linker was introduced via standard solid-phase DNA synthesis and unmasked to a 3′-thiol by DTT treatment. A 50 µL sample containing 5 nmol of DNA anchor oligonucleotide in 100 mM HEPES, pH 7.5, and 50 mM DTT was incubated at 37 °C for 2 h. The reduced product was precipitated to remove excess DTT by addition of 50 µL of water, 10 µL of 3 M NaCl and 300 µL of ethanol. The precipitated product (DNA-HEG-p-C_3-SH) was dissolved in 35 µL of water and 10 µL of 100 mM triethylammonium acetate, pH 7.0. Activation as the pyridyl disulfide was achieved by adding 5 µL of 100 mM 2,2′-dipyridyl disulfide in DMF and incubating at 37 °C for 2 h. The product (DNA-HEG-p-C_3-SSPy) was precipitated by addition of 50 µL of water, 10 µL of 3 M NaCl and 300 µL of ethanol and dissolved in 25 µL of water. Conjugation to the peptide was performed by adding 20 µL of 50 mM triethylammonium acetate, pH 7.0 and 5 µL of 20 mM peptide (100 nmol, 20 equivalents). The sample was incubated at 37 °C for 12 h, and the DNA-anchored peptide was purified by 20% PAGE. A typical yield was 3.5–4.5 nmol.

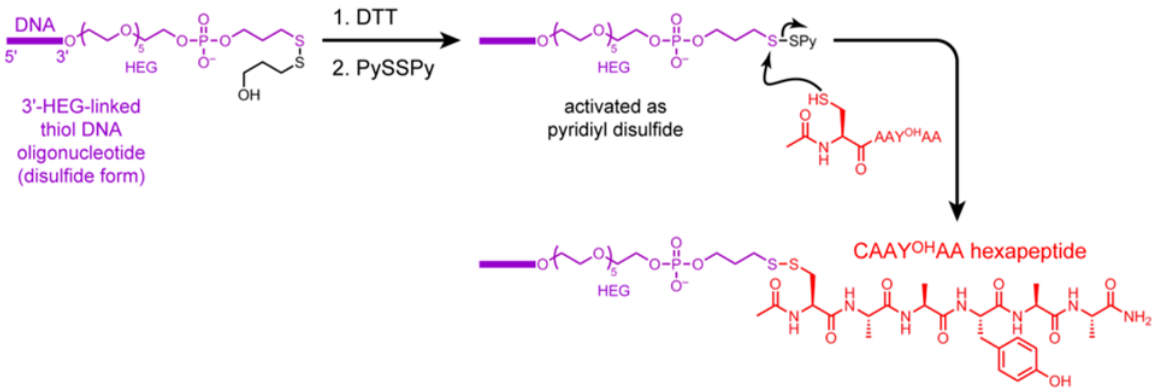

Figure 4.10: Synthesis and structure of the DNA-anchored CAAYAA hexapeptide phosphorylation substrate. Note the hexa(ethylene glycol), or HEG, tether connecting the DNA oligonucleotide anchor to the peptide.

4.4.2 In vitro selection procedure

For the tyrosine kinase selections, the PCR step was performed as described in Chapter 2, except that Pfu polymerase was used in place of Taq polymerase because the former led to substantially greater yield of the desired PCR product. T4 DNA ligase was used to ligate each pool molecule to the DNA anchor. For the JN1, JQ1 and JS1 selections, the splint sequence for the ligation step in all rounds was 5′-ATATGTCTTTCAATAGTCCCCATCAGGATCAGCTCTATTGAAAGACATAT-3′. Nucleotide details for the selection and capture steps of the JS1 selection are shown in Figure 4.11. For the other tyrosine kinase selections, the sequences to the 3′-side of the pool's random region (and of 8VP1 catalytic region) were complementary to the sequences of DNA anchors (shown in the previous section). Additionally, the sequence of the MB ATP aptamer used in the FK1-FM1 selections was 5′-CCACTGAGGGGCGAGGAGAG-AAACTGCGGGGAGTGG-3′ and the primer-binding sequences to the 5′-side of the random regions were 5′-CGAAGCGCTAGAACAT-3′ (FD1-FF1 selections), 5′-CGAAA-TGATGGCTATTTC-3′ (JN1 and JQ1 selections) and 5′-CGAAATAGATTATCATTC-3′ (JP1 selection). While for the modular selections the aptamer was used to bind ATP, in the direct (nonmodular) selections, the left-hand DNA binding arm (i.e., the pool segment to the 5′-side of the random region) could not interact with ATP and was also not base-paired to anything. The 8VP1 deoxyriozyme was validated to be highly selective for reaction of phosphorylated Tyr (Tyr^{P}) over unphosphorylated Tyr (Tyr^{OH}).[41] The sequence of the 8VP1 catalytic region was 5′-GGACACGATGAGTGACTAAGTGGAATGAGGAAAG-CACGAG-3′. To avoid emergence of noncatalytic DNA sequences that migrate aberrantly at a fixed PAGE position, the length of the capture oligo reaction partner was alternated in

successive rounds between a short version (17 nt pppRNA, odd-numbered rounds including round 1) and a long version (54 nt pppRNA-DNA chimera, even-numbered rounds). The pppRNA-DNA chimera was prepared by splint ligation of the pppRNA component to the 5′-phosphorylated DNA component using a complementary DNA splint and T4 DNA ligase. The short and long acceptor (capture) oligonucleotides had different sequences to avoid emergence of deoxyribozymes that catalyze unwanted covalent attachment of Tyr^{OH} to RNA during the capture step (this would be facilitated if the DNA random region had a sequence that allows substantial Watson-Crick base pairing to the constant capture oligo sequence). For the long capture oligo, the 8VP1 binding arm was extended several nucleotides past the RNA-DNA junction, which suppressed nonspecific cleavage by transesterification of the RNA nucleotides near the junction.

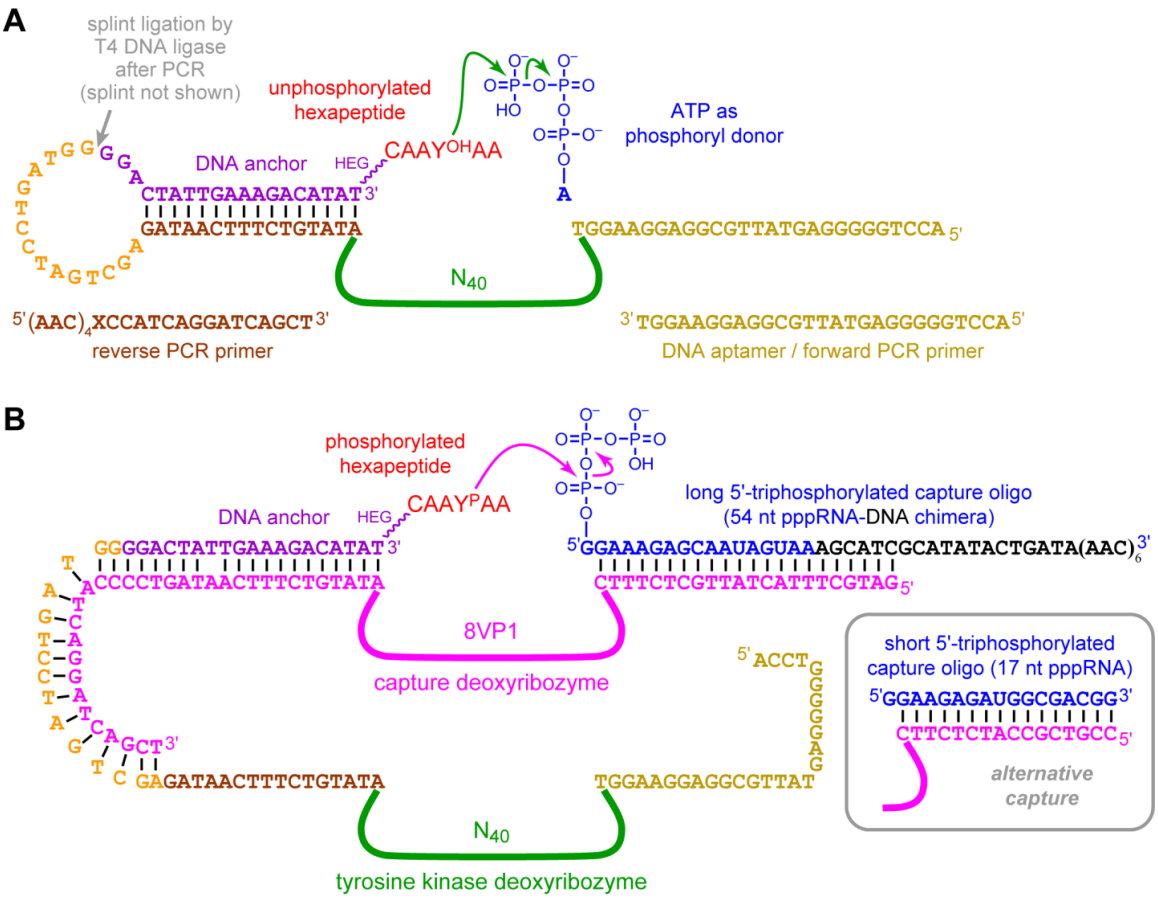

Figure 4.11: Nucleotide details of the selection and capture steps of the JS1 selection (other modular tyrosine kinase selection experiments were performed analogously). (A) Selection step. (B) Capture step.

4.4.3 MALDI mass spectrometry of the 14JS101 phosphorylation product

The DNA-anchored CAAY^{P}AA phosphorylation product was prepared from a 10 µL sample containing 50 pmol of DNA-anchored HEG-tethered CAAY^{OH}AA substrate (Fig. 4.12), 80 pmol of 14JS101 deoxyribozyme and 2 nmol of ATP, which were annealed in 5 mM Tris, pH 7.5, 15 mM NaCl and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed phosphorylation reaction was initiated by bringing the sample to 20 µL total volume containing 40 mM Tris, pH 7.5, 1.0 mM $ZnCl_2$, 10 mM $MnCl_2$, 40 mM $MgCl_2$ and 150 mM NaCl. The sample was incubated at 37 °C for

14 h, quenched with EDTA, desalted by Millipore C_{18} ZipTip and analyzed by MALDI mass spectrometry (Bruker UltrafleXtreme; matrix 3-hydroxypicolinic acid, positive ion mode).

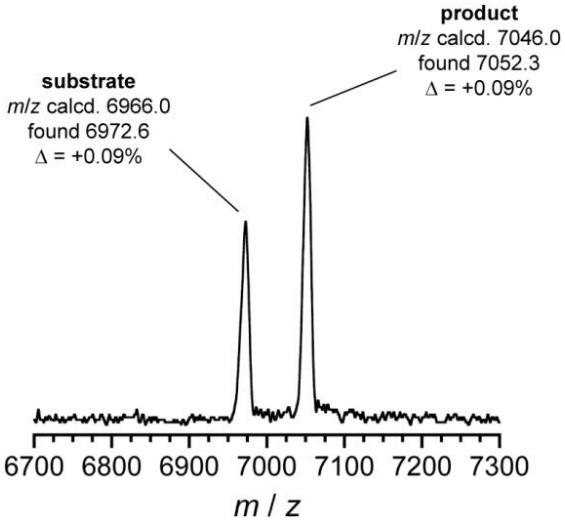

Figure 4.12: MALDI mass spectrometry of the 14JS101 phosphorylation product.

4.5 References

1. Wong, O. Y.; Pradeepkumar, P. I.; Silverman, S. K., DNA-catalyzed covalent modification of amino acid side chains in tethered and free peptide substrates. *Biochemistry* **2011,** *50*, 4741-4749.
2. Sachdeva, A.; Silverman, S. K., DNA-catalyzed serine side chain reactivity and selectivity. *Chem. Commun.* **2010,** *46*, 2215-2217.
3. Pradeepkumar, P. I.; Hobartner, C.; Baum, D. A.; Silverman, S. K., DNA-catalyzed formation of nucleopeptide linkages. *Angew. Chem. Int. Ed.* **2008,** *47*, 1753-1757.

4. Breaker, R. R.; Joyce, G. F., A DNA enzyme that cleaves RNA. *Chem. Biol.* **1994,** *1*, 223-229.
5. Santoro, S. W.; Joyce, G. F., A general purpose RNA-cleaving DNA enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **1997,** *94*, 4262-4266.
6. Silverman, S. K., In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic Acids Res.* **2005,** *33*, 6151-6163.
7. Schlosser, K.; Li, Y., A versatile endoribonuclease mimic made of DNA: characteristics and applications of the 8-17 RNA-cleaving DNAzyme. *ChemBioChem* **2010,** *11*, 866-879.
8. Chandra, M.; Sachdeva, A.; Silverman, S. K., DNA-catalyzed sequence-specific hydrolysis of DNA. *Nat. Chem. Biol.* **2009,** *5*, 718-720.
9. Li, Y.; Liu, Y.; Breaker, R. R., Capping DNA with DNA. *Biochemistry* **2000,** *39*, 3106-3114.
10. Li, Y.; Breaker, R. R., Phosphorylating DNA with DNA. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96*, 2746-2751.
11. Huizenga, D. E.; Szostak, J. W., A DNA aptamer that binds adenosine and ATP. *Biochemistry* **1995,** *34*, 656-665.
12. Barbu, M.; Stojanovic, M. N., A fresh look at adenosine-binding DNA motifs. *ChemBioChem* **2012,** *13*, 658-660.
13. Lin, C. H.; Patel, D. J., Structural basis of DNA folding and recognition in an AMP-DNA aptamer complex: distinct architectures but common recognition motifs for DNA and RNA aptamers complexed to AMP. *Chem. Biol.* **1997,** *4*, 817-832.

14. Jhaveri, S. D.; Kirby, R.; Conrad, R.; Maglott, E. J.; Bowser, M.; Kennedy, R. T.; Glick, G.; Ellington, A. D., Designed Signaling Aptamers that Transduce Molecular Recognition to Changes in Fluorescence Intensity. *J. Am. Chem. Soc.* **2000,** *122*, 2469-2473.
15. Stojanovic, M. N.; de Prada, P.; Landry, D. W., Fluorescent Sensors Based on Aptamer Self-Assembly. *J. Am. Chem. Soc.* **2000,** *122*, 11547-11548.
16. Stojanovic, M. N.; Kolpashchikov, D. M., Modular aptameric sensors. *J. Am. Chem. Soc.* **2004,** *126*, 9266-9270.
17. Nutiu, R.; Li, Y., In Vitro Selection of Structure-Switching Signaling Aptamers. *Angew. Chem. Int. Ed.* **2005,** *44*, 1061-1065.
18. Cho, E. J.; Yang, L.; Levy, M.; Ellington, A. D., Using a deoxyribozyme ligase and rolling circle amplification to detect a non-nucleic acid analyte, ATP. *J. Am. Chem. Soc.* **2005,** *127*, 2022-2023.
19. Liu, J.; Lu, Y., Fast Colorimetric Sensing of Adenosine and Cocaine Based on a General Sensor Design Involving Aptamers and Nanoparticles. *Angew. Chem. Int. Ed.* **2006,** *45*, 90-94.
20. Zuo, X.; Song, S.; Zhang, J.; Pan, D.; Wang, L.; Fan, C., A target-responsive electrochemical aptamer switch (TREAS) for reagentless detection of nanomolar ATP. *J. Am. Chem. Soc.* **2007,** *129*, 1042-1043.
21. Li, D.; Shlyahovsky, B.; Elbaz, J.; Willner, I., Amplified Analysis of Low-Molecular-Weight Substrates or Proteins by the Self-Assembly of DNAzyme-Aptamer Conjugates. *J. Am. Chem. Soc.* **2007,** *129*, 5804-5805.

22. Ali, M. M.; Li, Y., Colorimetric sensing by using allosteric-DNAzyme-coupled rolling circle amplification and a peptide nucleic acid-organic dye probe. *Angew. Chem. Int. Ed.* **2009,** *48*, 3512-3515.

23. Zuo, X.; Xiao, Y.; Plaxco, K. W., High specificity, electrochemical sandwich assays based on single aptamer sequences and suitable for the direct detection of small-molecule targets in blood and other complex matrices. *J. Am. Chem. Soc.* **2009,** *131*, 6944-6945.

24. Cho, E. J.; Lee, J. W.; Ellington, A. D., Applications of aptamers as sensors. *Annu. Rev. Anal. Chem.* **2009,** *2*, 241-264.

25. Wang, Y.; Li, Z.; Hu, D.; Lin, C. T.; Li, J.; Lin, Y., Aptamer/graphene oxide nanocomplex for in situ molecular probing in living cells. *J. Am. Chem. Soc.* **2010,** *132*, 9274-9276.

26. Gulbakan, B.; Yasun, E.; Shukoor, M. I.; Zhu, Z.; You, M.; Tan, X.; Sanchez, H.; Powell, D. H.; Dai, H.; Tan, W., A dual platform for selective analyte enrichment and ionization in mass spectrometry using aptamer-conjugated graphene oxide. *J. Am. Chem. Soc.* **2010,** *132*, 17408-17410.

27. Xiang, Y.; Lu, Y., Using personal glucose meters and functional DNA sensors to quantify a variety of analytical targets. *Nat. Chem.* **2011,** *3*, 697-703.

28. Alila, K. O.; Baum, D. A., Modulation of an RNA-branching deoxyribozyme by a small molecule. *Chem. Commun.* **2011,** *47*, 3227-3229.

29. Tang, Y.; Ge, B.; Sen, D.; Yu, H. Z., Functional DNA switches: rational design and electrochemical signaling. *Chem. Soc. Rev.* **2014,** *43*, 518-529.

30. Levy, M.; Ellington, A. D., ATP-dependent allosteric DNA enzymes. *Chem. Biol.* **2002,** *9*, 417-426.

31. Dirks, R. M.; Pierce, N. A., Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci. U.S.A.* **2004,** *101*, 15275-15278.

32. Wieland, M.; Benz, A.; Haar, J.; Halder, K.; Hartig, J. S., Small molecule-triggered assembly of DNA nanoarchitectures. *Chem. Commun.* **2010,** *46*, 1866-1868.

33. Seemann, I. T.; Singh, V.; Azarkh, M.; Drescher, M.; Hartig, J. S., Small-molecule-triggered manipulation of DNA three-way junctions. *J. Am. Chem. Soc.* **2011,** *133*, 4706-4709.

34. Murphy, F. L.; Cech, T. R., Alteration of substrate specificity for the endoribonucleolytic cleavage of RNA by the *Tetrahymena* ribozyme. *Proc. Natl. Acad. Sci. U.S.A.* **1989,** *86*, 9218-9222.

35. Khosla, C.; Harbury, P. B., Modular enzymes. *Nature* **2001,** *409*, 247-252.

36. Urata, H.; Nomura, K.; Wada, S.; Akagi, M., Fluorescent-labeled single-strand ATP aptamer DNA: chemo- and enantio-selectivity in sensing adenosine. *Biochem. Biophys. Res. Commun.* **2007,** *360*, 459-463.

37. Wang, B.; Huang, F.; Nguyen, T.; Xu, Y.; Lin, Q., Microcantilever-Based Label-Free Characterization of Temperature-Dependent Biomolecular Affinity Binding. *Sens. Actuators. B Chem.* **2013,** *176*, 653-659.

38. Walsh, S. M.; Sachdeva, A.; Silverman, S. K., DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc.* **2013,** *135*, 14928-14931.

39. Chandrasekar, J.; Silverman, S. K., Catalytic DNA with phosphatase activity. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110*, 5315-5320.

40. Velez, T. E.; Singh, J.; Xiao, Y.; Allen, E. C.; Wong, O. Y.; Chandra, M.; Kwon, S. C.; Silverman, S. K., Systematic evaluation of the dependence of deoxyribozyme catalysis on random region length. *ACS Comb. Sci.* **2012,** *14*, 680-687.
41. Sachdeva, A.; Chandra, M.; Chandrasekar, J.; Silverman, S. K., Covalent Tagging of Phosphorylated Peptides by Phosphate-Specific Deoxyribozymes. *ChemBioChem* **2012,** *13*, 654-657.
42. Lorsch, J. R.; Szostak, J. W., In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature* **1994,** *371*, 31-36.
43. Lorsch, J. R.; Szostak, J. W., Kinetic and thermodynamic characterization of the reaction catalyzed by a polynucleotide kinase ribozyme. *Biochemistry* **1995,** *34*, 15315-15327.
44. Wilson, C.; Szostak, J. W., In vitro evolution of a self-alkylating ribozyme. *Nature* **1995,** *374*, 777-782.
45. Nowakowski, J.; Shim, P. J.; Prasad, G. S.; Stout, C. D.; Joyce, G. F., Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme. *Nat. Struct. Biol.* **1999,** *6*, 151-156.